INVEST

INVEST Working Papers 71/2023

Better close to home? Geographical and socioeconomic constraints on gendered educational transitions at the upper secondary level

Irene Prix
Outi Sirniö
Juhani Saari

04.04.2023

ISSN 2737-0534

The Inequalities, Interventions, and New Welfare State (INVEST) ecosystem brings together the efforts of over 200 researchers and advances cross-disciplinary research. It forms an international collaboration platform that connects researchers, practitioners, public organisations, third sector organisations, companies, students, families and other citizens that work with us in our mission to increase the wellbeing of the Finnish society during childhood, youth and early adulthood. The INVEST ecosystem can be harnessed to solve welfare state and public sector problems and provide greater insights to governments as they form policy. It is involved in large-scale policy reforms. INVEST is funded by the Academy of Finland Flagship Programme and the host institutions and affiliates multiple projects funded by various external research funders. The INVEST working paper series has subseries for the long-term projects funded by the Strategic Research Council (SRC).

Better close to home?
Geographical and socioeconomic constraints on gendered educational transitions at the upper secondary level

Irene Prix, INVEST flagship research centre, University of Turku, Finland
Outi Sirniö, Finnish Institute for Health and Welfare, Finland
Juhani Saari, Statistics Finland, Finland

Corresponding author: Irene Prix, irepri@utu.fi

Abstract

Educational decisions are affected by geographical accessibility, which may have far-reaching consequences on future educational pathways. In this paper, we examine the extent to which geographical distance to educational institutions may moderate young people's applications to upper secondary education in terms of both the track and the gender-(a)typicality of vocational fields of study they apply to. Our study relies on rich register-based data of complete cohorts of 16-year-olds applying to Finnish upper secondary institutions, linked with geographical information on their closest educational alternatives. We find that travel time to high schools is more decisive than the distance to vocational schools, with geographical accessibility being more significant for boys' rather than for girls' application patterns. Moreover, distance sensitivity varied by social origin, with daughters of low-educated parents and sons of medium-educated parents particularly likely to adjust their upper secondary application to the geographical accessibility of educational alternatives. However, we find some indications that particularly girls from lower-educated social backgrounds are more prepared to consider fields of study not typical for their gender if they are more geographically accessible than key alternatives. No such gender-atypical substitutions were evident among boys. We discuss the implications of these findings in the context of explanatory approaches based on risk aversion and (gender) socialization.

Keywords: geographical distance; upper secondary education; educational track; field of study; social origin; gender

Declarations of interest: none

Acknowledgements

We thank Matti Koskela (Regional State Administrative Agency for Eastern Finland) for kindly providing this study with data on teaching location geocodes. Previous versions of this article have been presented at the ECSR annual conference (2018, 2021) and the RC28 Spring Meeting (2021), as well as in research seminars at the University of Turku, Finland (2021), the Finnish Institute for Health and Welfare (2021), and the University of Pavia, Italy (2022). We thank participants for their constructive feedback. All views and errors in this paper are the authors' own and should not be attributed to their affiliated organizations.

Funding

This research has been supported by the INVEST Research Flagship Center, funded by the Academy of Finland Flagship Programme [grant number: 320162]. Irene Prix acknowledges partial funding by the Turku Institute for Advanced Studies. Juhani Saari's contribution was partly funded by the Kone Foundation (PAHE-project, PI: Sonja Kosunen).

1. Introduction

Educational inequality has distinct geographical dimensions. Young people living at a longer geographical distance from higher education institutions have been found to participate less often in tertiary education (Finger, 2016; Frenette, 2006; Sá et al., 2006; Spiess & Wrohlich, 2010; Turley, 2009). Some studies have found distance to induce a shift among entrants towards more geographically accessible institutions rather than to affect tertiary enrolment rates as such (Flannery & Cullinan, 2014; Gibbons & Vignoles, 2012; Knutsen et al., 2020). Similarly, the distance to education seems to play some role in students' field of study choice (Hango et al., 2019; Suhonen, 2014).

While geographical accessibility may not on its own guarantee equal educational opportunities, existing evidence suggests that travel distances are relevant for educational decisions, particularly among young people from lower socioeconomic backgrounds (Denzler & Wolter, 2011; Gibbons & Vignoles, 2012; Hällsten, 2010; Pigini & Staffolani, 2016). If geographical accessibility affects young adults' educational pathways, longer travel times likely pose even greater constraints, understood in terms of social, psychological and financial costs, on even younger students, such as adolescents considering their educational options. Particularly in tracked educational systems, geographical inequality at the point of upper secondary transitions may thus have far-reaching repercussions on young people's future pathways. With a handful of exceptions (Bertrand-Cloodt et al., 2010; Dickerson & McIntosh, 2013; Falch et al., 2013; Virtanen & Riukula, 2021), upper secondary education has been fairly neglected by (quantitative) sociological inquiries into the geographical dimensions of educational inequality.

In this study, we analyse vertical and horizontal dimensions of upper secondary transitions in Finland by combining information on educational applications and geographical distance to educational institutions. In Finland, the choice between academic and vocational tracks is the first turning point with "vertical" implications regarding educational attainment, as high schools represent a more direct and more likely pathway towards tertiary enrolment. Geographical constraints at this first significant educational junction may thus not only constitute an incidence of educational inequality in its own right, but significantly exacerbate existing social origin differences in educational outcomes.

Upper secondary transitions are not only relevant to intergenerational social mobility or reproduction processes, but they also are highly gendered. Higher education qualifications in Finland are significantly more common among women than men (Statistics Finland, 2018). This gender difference in average educational attainment is partly rooted in the fact that girls are now overrepresented on the high school track, whereas boys constitute the majority among students on the vocational track of upper secondary education. In addition, also field specializations within the vocational sector are highly gender-segregated. As such, upper secondary education constitutes a complex junction where social reproduction and mobility processes combine with gender dynamics.

Our aim in this paper is to examine how geographical constraints may interact with both social inequality and gender segregation processes in shaping adolescents' educational aspirations in Finland. In doing so, we offer both empirical and theoretical contributions. Firstly, as pointed out above, aspects of geographical accessibility constitute a neglected angle in social research on educational opportunity, particularly with regard to upper secondary education. Secondly, our

analyses of geographical constraints in education open up a new angle on the interaction between social origins and gender segregation processes in education (Prix & Kilpi-Jakonen, 2022; Seehuus, 2019). Distance to education may aggravate tensions for young people between aspired and feasible educational pathways, especially when geographically accessible options vary with regard to their social mobility prospects and the degree to which they would conform or oppose gender normative expectations.

In summary, these aims lead us to the following two research questions:

1. How does geographical distance matter for adolescents' decisions between vocational and academic tracks on their upper secondary applications, and to what extent does its relevance vary by social origin?

2. To what extent may the relative proximity of educational alternatives increase shifts in preferred tracks and fields of study, particularly with regard to gender-(a)typical substitutions, and how do such substitutions vary by social origin?

We examine these questions using rich register-based data on educational applications of complete cohorts of 16-year-old girls and boys applying to the Finnish upper secondary sector between 2009 and 2013. The fact that this paper is based on young people's applications rather than their enrolments constitutes a third aspect of our contribution, given that it allows us to observe a variety of educational pathways that the young applicants have considered (as their first and next-best choice) before any institutional selection processes further limit their options. Our aim to consider substitution patterns poses some methodological challenges, which we tackle in this article by introducing mother logit models to the sociological research context.

2. Background and hypotheses

2.1 Why and how should geographical distance matter for upper secondary track choices?

Financially speaking, longer distances to education typically involve higher travel costs. If commuting is not feasible, young people may have to move out of the parental home into student accommodation closer to the school. This raises expenses due to relocation costs, rent and maintaining a separate household. At the same time, the time spent commuting or living away from local friends and family may lead to also significant emotional costs. Qualitative research has highlighted that young people in rural and sparsely populated areas are often strongly attached to their local environment, with social ties centring significantly on family and (often locally residing) relatives (Rönnlund, 2020; Tuuva-Hongisto et al., 2016). As such, longer distances may weaken social ties to adolescents' home environment and increase pressures to become independent at an early age. For students enrolled in education far from home, these overall costs of distance may accumulate over time, which may explain the greater risk of drop-out observed among adolescents studying further away from home (Falch et al., 2013; Virtanen & Riukula, 2021). Although perceptions of distance may vary depending on local contexts, infrastructure, and reference points (Pulkkinen & Rautopuro, 2018, p. 19), previous research suggests that on average, young people prefer educational alternatives in relative proximity over such located further away (Dickerson & McIntosh, 2013; Gibbons & Vignoles, 2012; Knutsen et al., 2020). Based on these considerations and prior findings, we expect the following:

Expectation 1: On average, the longer the distance between applicants' home and either upper secondary track, the lower the probability that the applicant ranks this track as their first choice.

At the same time, it is plausible that gender may play a role in how sensitive young people are to the geographical distance to education, given the significant differences between men's and women's average educational pathways observed in Finland as well as internationally. Some scholars have argued that educational opportunities and occupational domains linked to female-dominated public sector and service occupations tend to be more strongly linked to urbanized regions, while male-dominated domains are better represented also in rural areas (Bock, 2017). Realizing their (gender-normative) educational and occupational aims may thus require on average greater compromises regarding their geographical location from young women compared to young men (Armila et al., 2018; Faggian et al., 2007).[1]

Prior evidence both supports and contradicts this expectation of distance mattering less for young women than men, suggesting that variation in culture and opportunity structure may affect gender differences in distance sensitivity. A recent study of upper secondary school closures in Finland found boys to react more strongly than girls to the closure of local alternatives (Virtanen & Riukula, 2021). The reverse gender pattern has however been reported in England, in the only other study on the distance sensitivity of adolescents' upper secondary track choices we are aware of (Dickerson & McIntosh, 2013). Research focusing on tertiary education has similarly found mixed results, ranging from no noteworthy gender differences (Gibbons & Vignoles, 2012; Suhonen, 2014) to higher distance sensitivity among young women rather than men (Frenette, 2006).

Based on prior Finnish findings on upper secondary distances, we expect the following gender differences:

Expectation 2: Girls are on average less sensitive to the distance of upper secondary institutions than boys when selecting their first-ranked track on their upper secondary application.

2.2 Distance sensitivity and social inequality in education: culture, risk and uncertainty

If educational pathways are psychologically and financially more costly the more distant they are, the accessibility of these pathways will depend on whether applicants are able and willing to afford these costs. As such, spatial dimensions of education are likely to feed into larger social inequality dynamics playing out within the educational system.

Cultural class accounts assume that children from less educated backgrounds forgo academic tracks disproportionately often due to the educational system, which, dominated by cultural elites, is systematically biased against the working class (e.g., Bourdieu & Passeron, 1990). From this perspective, it is not immediately clear that geographical distance or proximity should matter at all. As the cultural class account assumes cultural alienation from academic tracks as the primary motor

[1] Although high schools are on average more geographically accessible than vocational programmes, female-typical health and social care fields tend to be located in greater distance than engineering-related vocational programmes in Finland (Table 1).

for social origin differentials in education, this implies that particularly children from less educated backgrounds should be fairly insensitive to local variations in educational opportunities.

Relative risk aversion approaches, by contrast, assume that families in all social classes share the goal of avoiding social downward mobility for their children, whereas possibilities for upward mobility are only of secondary importance, albeit not irrelevant (Goldthorpe, 2007). For families in advantaged social positions, this means that educational pathways not clearly associated with higher education destinations may be considered risky with regard to this goal of status maintenance. By contrast, for working-class families, academic pathways carry more uncertainty than the more familiar and predictable vocational trajectories. This is because skilled working-class parents may have more resources and knowledge to support their children throughout vocational school compared to high school education. This should minimize the relative risk of vocational school failure and the loss of time and resources linked to switching to vocational tracks if the attempt of studying on academic tracks should turn out to be unsuccessful (Goldthorpe, 1996, 2007).

Recently, some of the central elements of the risk aversion model have come under closer scrutiny as well as criticism (Barone et al., 2021; Hällsten, 2017), mainly regarding the assumptions that are implied by the formal representation of the risk aversion model (Breen & Goldthorpe, 1997). In this article, we interpret the risk attached to educational pathways not as (subjective) calculable probabilities of longer-term failure that parents minimize in a sophisticated calculus, but rather as *uncertainty* that is difficult to quantify, to which parents are assumed to respond with fairly simple heuristics in weighing their options (e.g., Kozyreva & Hertwig, 2021).

Based on this interpretation of the risk aversion perspective, sensitivity to the geographical distance of upper secondary options is likely to vary across families in different socioeconomic locations. First, we assume that families are least sensitive to the distance of those educational options that represent the most predictable pathway for their children to reproduce their parents' socioeconomic position. In this sense, distance to a status-reproducing type of education may appear less costly, whereas the distance to more unfamiliar options may compound uncertainties about the consequences of possible failure and their longer-term implications. On the other hand, although parents with shorter educational qualifications may be overall more reserved towards the academic pathway, a relative decrease in the distance to high school may make this option (and its upward mobility prospects) slightly less costly, which may also lower uncertainty differentials vis-a-vis the fall-back option of vocational pathways. On the other hand, distance to upper secondary pathways associated with educational downward mobility, such as the vocational option in the case of children with highly educated parents, should not modify the risks associated with it. In summary, we expect the following patterns:

Expectation 3. Sensitivity to distance varies by social background and upper secondary track

a. Children from lower- and medium-educated social origins should be more sensitive to the distance of high schools compared to their peers with highly educated parents when applying to upper-secondary education.
b. For applicants from all social origins, distance to the nearest vocational school should be less relevant than high school distances for deciding on the first-ranked upper secondary track.

2.3 Compromises between socioeconomic and cultural pressures? Social mobility prospects and gender segregation in upper secondary education

Given that vocational school branches in Finland can differ in the type of programmes they offer, the geographical distance to vocational schooling can vary noticeably between vocational fields. This also means that geographical distance may moderate the accessibility of upper secondary education not only with regard to educational tracks but also in terms of vocational fields of study, particularly for applicants with lower resources. Is it likely that the relative geographical proximity of particular educational options could lead young people to make pragmatic compromises, even if this would imply crossing into gender-atypical fields?

Some educational psychologists have argued that children's occupational interests develop through a process of excluding unsuitable and inaccessible options, leading to a formation of a gradient of tolerable options rather than a discrete set of expressive preferences. In this account, most children tend to consider occupations through their expressive interests only after first eliminating vocational alternatives deemed unsuitable with respect to their occupational sex type and associated social status. As a consequence, this model expects that gender-atypical alternatives will be considered only as a very last resort, with young people being instead more ready to compromise on their expressive interests and the prestige level of their aspired occupation (Gottfredson & Lapan, 1997).

But scholars have argued that the strength of gender stereotypes mediated via socialization may however vary by social origin, assuming more restrictive gender socialization as fostering sex-type conformity, particularly in families with lower resources (Chesters, 2021). Social control theory, on the other hand, has countered that socialization on its own does not suffice to maintain gender-segregated structures, suggesting that compliance with gendered boundaries depends instead on the strength of social sanctions and the opportunities to cross into gender-atypical domains (Jacobs, 1989).

While research on young people's field of study preferences has highlighted the importance of expressive rather than rational motivations (Barone & Assirelli, 2020), the question remains as to what structural opportunities may lead adolescents to contemplate pragmatic compromises. A Dutch study of vocational entrants (which was however not restricted to young people nor differentiated by social origin) found geographical distance to more strongly moderate women's rather than men's take-up of (male-dominated) engineering courses (Bertrand-Cloodt et al., 2010). Recent research has found support for sex-type compromises to depend on social mobility outlooks, contrasting the idea of gender-stereotypical orientations as a defining characteristic of less affluent social origins (Prix & Kilpi-Jakonen, 2022).

From the discussion above, we conclude that pragmatic compromises leading to gender-atypical application patterns may be most likely for those young people for whom geographical distance constitutes a significant constraint. By contrast, children from higher educated families, for whom vocational routes are in general less common, may also be less constrained in terms of resources and therefore experience lower pressures to compromise on their interests. Even if they display greater openness towards gender-atypical fields due to variation in socialization patterns, this preference should be less sensitive to distance compared to young people from socioeconomically less advantaged backgrounds. These considerations lead us to the following expectation:

Expectation 4: Compared to children from other social backgrounds, children of low-educated parents are more likely to pick a gender-atypical field as their first (or next-ranked) choice on their application if this field is geographically closer than other alternatives.

3. Context: Upper secondary education and geographical distances in Finland

For the cohort of adolescents in this paper, compulsory education ended with completing nine grades of comprehensive school. About 85% of compulsory school leavers enter upper secondary right away, with about 95% of the age cohort having enrolled by the time they were 17 years old (Statistics Finland, 2023). The upper secondary transition constitutes the first branching point in the Finnish educational system, with high schools as the traditional route towards tertiary education and vocational schools offering fairly occupation-specific qualifications facilitating labour market entry into skilled jobs.

Vocational programmes in Finland have nationally standardized curricula. In 2015, around 50 different vocational qualifications were on offer for young people applying to the upper secondary system (Cedefop, 2015). Vocational institutions are mandated by law to cooperate with regional businesses and to offer courses aligning with local labour market demands. Study programmes for smaller and more specialized fields tend to be more geographically scattered, while fields like engineering and healthcare are more widely available. Thus the allocation of places of study of different fields in different parts of the country reflects policy decisions based on both local labour market demand and various regional and educational policy constraints (Stenström & Virolainen, 2014)

Although both high school diplomas and vocational qualifications grant eligibility to apply to higher education, tertiary transitions are nevertheless rare for vocational graduates. In 2019, about 1% of vocational graduates continued at university within a year of completing upper secondary education, compared to 30% of high school graduates. Entering polytechnics, which offer more vocational-oriented bachelor's and master's degrees compared to universities, represents a somewhat more common path for vocational graduates, although high school graduates outnumber them also on this pathway. In 2019, around 12% of young people with a vocational qualification (compared to 20% of high school graduates) entered polytechnics within a year of completing upper secondary education (Vipunen Education Statistics Finland, 2022).

Before entering upper secondary education, young people apply to a centralized admission system, for which they can select up to five programmes (high schools or specialized vocational programmes) in order of preference. Admission to the upper secondary level is based mainly on the grade point average of the basic education certificate awarded at the end of the comprehensive school. The admission system proceeds in the preference order that applicants have indicated on their application form. This means that once an applicant is selected for a given programme, all lower-ranked choices on their application are automatically cancelled. In 2014, about 84% of 16-year-old applicants were admitted to their first choice (Vipunen Education Statistics Finland, 2023).

Fields of study are highly gendered in Finland, particularly so at the upper secondary level (Statistics Finland, 2018, pp. 30–31). Female students are typically overrepresented in programmes associated with the arts, health/social care and service fields, whereas male students are concentrated in programmes related to science, IT and technology. Although male-dominated vocational fields tend

to be associated with higher pay, they tend to be associated with on average higher unemployment risks compared to many female-dominated vocational programmes (Prix & Kilpi-Jakonen, 2022, tbls. 2–3).

4. Data, variables and methods

Our starting point in this paper is a full-population data set compiled from administrative registers (Statistics Finland, 2023), comprising all children born in Finland in 1993-1997 and living in mainland Finland during the year they turn 15 (n=306,234). We link these children with information on their parents and their household for the preceding years when they were of compulsory school age (ages 7-15). Children not living with at least one biological or foster/adoptive parent at age 15 are excluded (n=6151 or 2% of the data). Our analyses are restricted to those young people in the cohort who apply to upper secondary education during the year they turn 16 (91% of the cohort), which represents the normative point of transition in the Finnish upper secondary system. After listwise deleting observations with missing values on our independent variables (n=392 or 0.1%), our final analytical data set comprises 272,847 children.

To determine the geographical distance of available educational alternatives, we rely on official information on available programmes and their teaching locations published by the Finnish Agency for Education from 2009 to 2013. The Regional State Administrative Agency for Eastern Finland compiled this information into a database for internal research purposes (Koskela, 2020), which they have kindly made available for our study.

4.1 Dependent variables

Our first set of analyses focuses on young people's first-ranked *upper secondary track* (high school or vocational) on their application. Further models differentiate this outcome variable to include also the particular *field of study* if a vocational programme is ranked first, divided according to six broad fields of study based on an aggregated classification standard used by the Finnish National Agency for Education. Combining track and field information, the resulting variable can take on seven values: general education (=high school track), arts (incl. humanities and education), business and administration, engineering (incl. IT and natural sciences), agriculture, health and social care, and hospitality.

Most of our analyses focus on applicants' *first-ranked option*, but supplemental analyses extend the view also to the *next-best upper secondary programme* that young people considered. As applicants may use several ranks in their preference order for the same type of programme (e.g., different high schools, or the same type of vocational programme in different locations), this next-best choice (which excludes the first-ranked option from the choice set) does not necessarily strictly coincide with the second rank. Rather, it represents the first alternative in terms of track or field that applicants considered aside from their first choice.

4.2 Key independent variables

We measure *geographical distance* as the driving time from young people's home postcode (at the end of the year when they turn 15) to the exact address of the nearest upper secondary institution, using OpenStreetMap data via the osrmtime-package (version 1.3.3) for Stata (Huber & Rust, 2016). This distance is calculated separately for the nearest vocational school and the nearest high school.

For our analyses that differentiate upper secondary choices by (vocational) fields, we separately calculate the driving time to the closest high school and each of the six broad vocational fields outlined above. In all our models, driving time is measured in ten-minute units. Table 1 describes applicants' average geographical distance to upper secondary tracks and broad vocational fields.

Our operationalization of social background relies mainly on *parents' education*, using as our reference the parent with the highest educational attainment level when the child was aged 15. We differentiate between basic education, upper secondary qualifications, short tertiary qualifications (incl. bachelor degrees and other post-secondary qualifications), and master's (incl. higher) degrees. We focus on parental education as our key social origin variable, as previous research has highlighted resource-specific associations between family background and children's outcomes (Thaning, 2021). To further accommodate the multidimensionality of social origin, however, we also add children's equivalized *household income* to the model, which we averaged over their school years (age 7-15) and measured as percentile ranks.

Table 1. Driving distances to the nearest upper secondary branches, by track and vocational field (2009-2013).

Upper secondary track	Mean distance	Standard deviation	Lower quartile	Upper quartile
High school	11.0	9.8	4.8	13.8
Vocational school	19.6	18.2	7.8	26.3
Hum/Culture	21.7	21.5	9.2	27.2
Business	20.6	18.1	8.9	27.3
Engineering/IT	15.9	14.7	6.5	19.9
Agriculture/Nature	30.8	20.7	16.4	39.4
Health/Social	19.7	17.8	8.4	26.3
Hospitality	17.7	15.8	7.6	23.1

4.3 Controls

Including applicants' *grade point average* (GPA) achieved at the end of comprehensive education helps us separate out such social origin associations that derive from applicants adjusting their plans to their predicted admission chances. We control for children's *native language*, given the previously observed higher average educational attainment levels observed among the Swedish-speaking minority in Finland (Härtull & Saarela, 2021). Adding *migration background* aims to accommodate the higher average educational aspirations observed among children of immigrants in Finland (Kilpi-Jakonen, 2011). We also include the *family type* children lived in at age 15, as it may have socialization consequences that affect children's educational aspirations.

As absolute geographical distances vary throughout the country as well as between rural and more urban municipalities, we include controls for the (NUTS2-classified) *region in Finland* as well as the degree of *urbanization* of applicants' home municipality at the end of the year in which applicants turned 15. To take into account that educational options may not vary only in distance, but also in their regional labour market prospects, our field-specific models also control for *field-specific unemployment rates.* This measure is calculated for each application year, based on 25-54-year-olds with an upper secondary qualification in the sub-regional unit to which applicants' home municipality belongs.

Table 2 provides an overview of the geographical variation of distances to upper secondary education for our data. For an overview of all variables used in our analyses, please refer to Table 3.

Table 2. Driving times (in minutes) to the nearest upper secondary branches (2009-2013), by region and upper secondary track.

	High school distances				Vocational school distances			
Region	Mean	Std. dev.	Lower quartile	Upper quartile	Mean	Std. dev.	Lower quartile	Upper quartile
Helsinki/Uusimaa	6.8	5.4	3.4	8.2	8.1	6.0	4.2	10.1
South	11.2	10.6	5.4	13.9	13.9	12.6	6.3	17.7
West	11.8	8.0	6.6	14.9	15	11.7	6.8	20.6
North/East	14.4	12.4	6.1	18.5	18.6	16.9	7.1	26.6
Total	11.0	9.8	4.8	13.8	13.8	12.9	5.4	17.2

4.4 Method

Our first research question focuses on the choice between high schools and vocational schools, which is why our first set of analyses centres on *binary logit models*. We present these results in terms of average marginal effects, as we consider probabilities as more intuitive and relevant for interpreting the results than the (log) odds scale (see, e.g., Mize, 2019).

Once we extend our analyses to differentiate between track and vocational fields of study (research question 2), we draw on *mother logit models.* Similarly to conditional logit analyses, these models accommodate both individual-specific variables (i.e., characteristics of applicants) and alternative-specific variables (i.e., characteristics of the outcome alternatives). However, mother logit models additionally include cross-effects, which means that the choice of a given outcome alternative can depend also on the characteristics of competing alternatives. Due to the specification of such cross-effects, mother logit models (unlike multinomial and conditional logit models) circumvent the independence of irrelevant alternatives -assumption (IIA), which facilitates the use of these models for examining substitution patterns (Steenburgh, 2008).

Our mother logit models use a very flexible specification, where distance associations are allowed to vary between different tracks/fields as well as by parental education (modelled as the three-way interaction distance×alternative×parental education). We interpret the results based on average predicted probabilities, where we compare different field/track-specific distance constellations of three key educational alternatives: high schools, engineering, as well as health and social welfare. As absolute distances vary between regions, we take the closest locally observed distance to the nearest high school as the locally "closest" distance for our counterfactual scenarios. In the reference scenario, all three key alternatives are set to this distance. In scenarios where two of the key fields are defined as further away, we set their distance as being located 15 minutes further away than the two other key fields. The distance of the relatively "closer" option remains set to that of the nearest high school. Distances to other programmes than the key three fields are kept at their observed values in all scenarios.

Note that not all of these relative distance constellations are equally common empirically. For instance, in our data, high schools are very rarely more distant than health and engineering

programmes. Furthermore, it is fairly common that health programmes are located at a longer distance than engineering, while the reverse is very rare.

Although mother logit models have been discussed since the 1970s, they have only been sparsely used in economics (and to our knowledge, not at all in sociology), perhaps because they may produce results that do not align with random utility modelling (RUM) expectations (Hess et al., 2018; Steenburgh, 2008). However, given that RUM theory has itself been criticized for a lack of realism (reviewed in Hess et al., 2018), and does not in its classic form constitute the theoretical point of departure for this article, we argue that this feature of the mother logit model does not impede our analytical purposes.

As our analyses are based on full population rather than sample data, traditional statistical inference plays a lesser role in this paper. Although we nevertheless report standard errors and p-values, we interpret the results based on substantive effect sizes rather than statistical significance.

Table 3. Descriptives of variables used in the analyses.

A. Frequency distributions by gender (column %)

Categorical variables	Girls	Boys	Total
First-ranked upper secondary track			
High school	61.3	46.7	54.0
Vocational school	38.7	53.3	46.0
First-ranked broad field			
General (=high school track)	61.3	46.7	54.0
Hum/Culture	4.1	2.0	3.0
Business	5.2	4.5	4.8
Engineering/IT	5.3	40.3	22.8
Agriculture/Nature	2.1	1.9	2.0
Health/Social	15.3	1.7	8.5
Hospitality	6.6	2.9	4.8
Parents' highest level of education			
basic/unkown	3.6	3.4	3.5
upper secondary	37.8	37.1	37.5
short tertiary	39.5	39.7	39.6
master's or higher	19.1	19.7	19.4
Family type			
lives with both parents	66.0	67.1	66.5
single parent	21.9	21.0	21.4
stepfamily	12.2	11.9	12.1
Both parents are immigrants			
yes	1.2	1.1	1.2
Home municipality's urbanisation			
urban	64.3	64.3	64.3
semi-urban	18.9	19.1	19.0

rural	16.8	16.6	16.7
NUTS region			
Helsinki/Uusimaa	27.1	27.3	27.2
South	21.2	21.0	21.1
West	25.6	25.8	25.7
North/East	26.1	25.9	26.0
Native tongue			
Finnish	93.4	93.6	93.5
Swedish	5.3	5.3	5.3
Other	1.2	1.1	1.2

B. Means and standard deviations (SD) by gender

	Girls		Boys		Total	
Continuous variables	Mean	SD	Mean	SD	Mean	SD
Shortest high school distance (minutes)	11.0	9.8	11.0	9.8	11.0	9.8
Shortest vocational school distance (minutes)	13.8	12.9	13.8	12.9	13.8	12.9
Average equivalized household income (euro)	25,712.9	21,748.8	26,007.8	24,774.2	25,860.3	23,310.3
GPA in lower secondary school (range: 4-10)	8.0	1.0	7.5	1.1	7.7	1.1
Number of observations	136,496		136,351		272,847	

5. Results

5.1 How does geographical distance matter for the choice between vocational and high school tracks?

Examining how geographical distance is on average reflected in young people's application patterns to upper secondary education, we first estimated a main effects binary logit model for the choice between high schools and vocational schools (Table 4). With increasing travel distance between their home and the nearest high school, applicants appear to be on average less likely to select high schools as their first choice on their upper secondary application, net of other controls in the model (in line with expectation 1). For each ten-minute driving time to the nearest high school, our models estimate the probability of applying to this track to decrease by on average 2 percentage points (pp.) in the case of girls and 3 pp. for boys. This slightly higher sensitivity to geographical distance among boys compared to girls replicates previous research and conforms to our expectations (expectation 2). Note, however, that not all distances are equally relevant. For both girls and boys applying to upper secondary education, longer distances to vocational school only barely (by less than half a percentage point for every ten minutes driving) increase applicants' interest in high schools, whereas

a greater distance to the nearest high school more clearly deters some applicants from selecting high schools as their first-ranked choice.

To examine whether distance sensitivity depends on social origin, we next model the interaction between parental education and the shortest driving times to each upper secondary track. Among girls (Figure 1, left panel), distance sensitivity displays a clear social gradient (in line with expectation 3a). The geographical distance between home and high school appears to matter least for girls with highly educated parents (-1.4 pp for each ten-minute increase in driving time). Daughters of low-educated parents, on the other hand, appear more sensitive to the geographical accessibility of high schools, with each ten-minute increase in driving time lowering their application probability by 3.5 pp.

Among applicants from all social origin groups, vocational distances appear as less important for upper secondary track decisions than driving times to the nearest high schools (in line with expectation 3b). For girls with low-educated parents, vocational distances do contribute to the decision to select high schools as their first-ranked track, with increasing vocational distances making high school plans more likely. For the daughters of medium- and highly-educated parents, on the other hand, vocational distances appear to play no substantive role in applying to upper secondary education (Figure 1, left panel).

Table 4. Binary logit model of high schools as the first-ranked track on young people's upper secondary application. Average marginal effects (standard errors in parentheses).

	Girls			Boys		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Parents' education (ref. upper secondary)						
basic	-0.038***		-0.039***	-0.025***		-0.026***
	(0.006)		(0.006)	(0.007)		(0.006)
short tertiary	0.069***		0.068***	0.065***		0.064***
	(0.002)		(0.002)	(0.002)		(0.002)
master's or higher	0.158***		0.157***	0.159***		0.157***
	(0.004)		(0.004)	(0.003)		(0.003)
High school distance (1=10 min.)		-0.022***	-0.021***		-0.031***	-0.031***
		(0.002)	(0.002)		(0.002)	(0.002)
Vocational school distance (1=10min.)		0.004***	0.006***		0.010***	0.011***
		(0.001)	(0.001)		(0.001)	(0.001)

McFadden's pseudo-R^2	0.370	0.359	0.371	0.438	0.427	0.440
AIC	114,865.1	116,829.0	114,617.6	105,932.9	107,966.5	105,478.8
Observations	136,496	136,496	136,496	136,351	136,351	136,351

Notes: All models include controls for region, urbanization degree of home municipality, GPA in lower secondary school (squared), family type, average equivalized household income (squared), mother tongue, having at least one immigrant parent. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

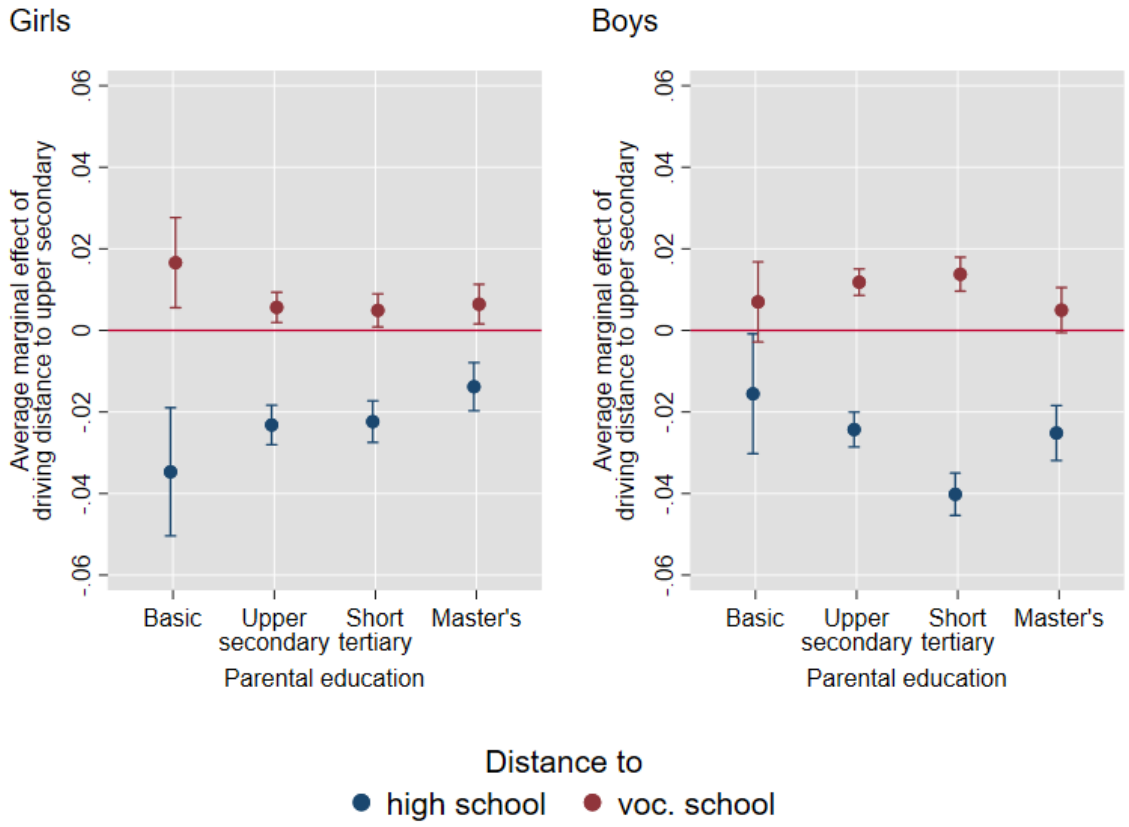


Figure 1. Distance sensitivity with regard to the top-ranked upper secondary track by parental background. Average marginal effects based on binary logit models with parental education×high school distance and parental education×vocational school distance -interactions.

Distance sensitivity patterns deviate a bit more strongly from our expectations among boys (Figure 1, right panel). Surprisingly, high school distances appear to play a similar role in deciding on the first-ranked track among the sons of highly-educated parents as among their peers whose parents have upper secondary or even lower qualifications (against expectation 3a). For these groups, longer driving times to high schools decrease the probability of selecting this track as their first choice by a

similar extent (about 2 pp. for each ten-minute increase). The strongest reaction to high school distances, on the other hand, is observed among the sons of parents with short-tertiary qualifications (-4pp for each ten-minute driving time increase). Among boys from all social backgrounds, driving times to vocational school matter less than those to high schools for their ranking of upper secondary tracks (in line with expectation 3b). However, sons of medium-level educated parents are somewhat more likely to consider high schools with increasing distance of vocational alternatives (about 1 pp. for each 10-minute increase in driving time to the nearest vocational school).

5.2 Compromises between tracks and (gender-typed) fields of study by social origin

Next, we examine to what extent applicants may consider high schools, (male-dominated) engineering, and (female-dominated) health and social welfare fields as possible substitutes if their relative proximity differs. To do so, we rely on mother logit models that include all cross-effects and three-way interactions between alternatives, distances and parental education.[2] We interpret results comparing model-based average predicted probabilities between several (hypothetical) scenarios, for which we manipulate the distance of three key alternatives: high schools, engineering, and the health and social welfare field. The reference scenario describes a situation where all three options are equidistant, using the observed distance of high schools as the reference distance for all three fields (leftmost column in Figure 2 and Figure 3, respectively).

[2] Model fit for our mother logit models surpassed that of equivalently specified conditional logit models in terms of both likelihood ratio tests and the Akaike Information Criterion (AIC). Model fit for mother logit models with differing interaction specifications (two-way vs. three-way) varied marginally, but remained very close overall (see Appendix Table A.1).

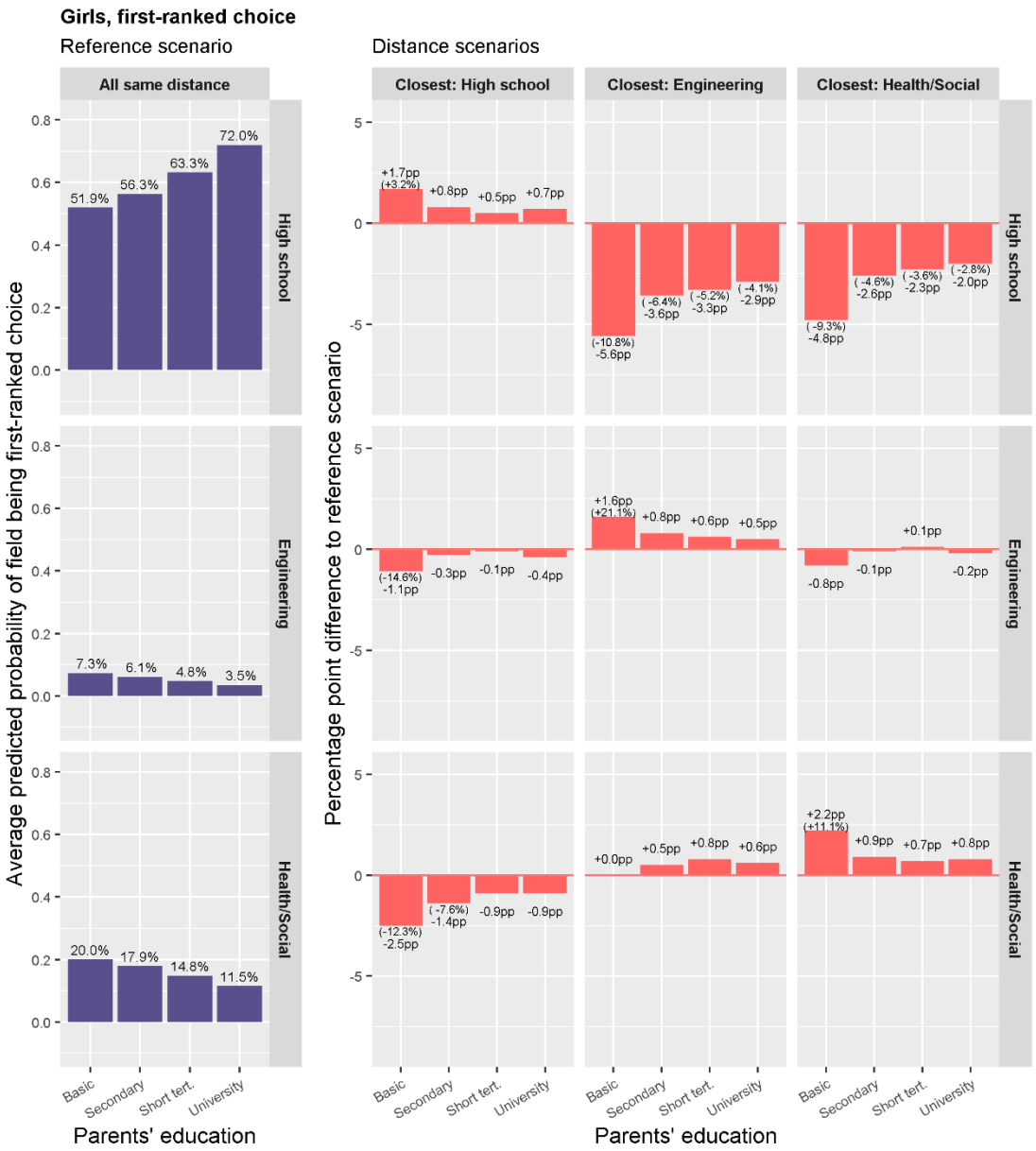


Figure 2. Girls' first-ranked choice by parental education and educational distance scenario. Average predicted probabilities in three hypothetical scenarios compared to a reference scenario (leftmost column). Results based on mother logit models with a three-way distance×alternative×parental education interaction. For included predictors and goodness of fit statistics, see Appendix Table 1 (MoL model 2, girls).

Among girls, a first-ranked high school application is 20 pp more likely for girls with highly-educated rather than low-educated parents, if key upper secondary alternatives are equidistant (Figure 2, reference scenario, top row). In a scenario where high schools are in relatively closer proximity than health or engineering, the model predicts an increase in first-ranked high school applications particularly among girls with low-educated (Figure 2, second column from left, top row). Conversely, while the absence of nearby high schools lowers first-ranked high school applications across the board (Figure 2, third and last column from left, top row), the predicted reduction is particularly pronounced among these girls with low parental education (by about 5 pp, which corresponds to a reduction of about 10-11%). This further exacerbates the clear social gradients in first-ranked high school applications observed in the reference scenario.

In a scenario where (male-dominated) engineering is relatively more accessible than other key alternatives, our model predicts a greater increase in engineering applications among girls with low-educated parents compared to their peers from other social origins (Figure 2, third column from left, middle row). By the same token, the results suggest that girls with low-educated parents are similarly more open to (gender-typical) health and social welfare programmes if this alternative is the most proximate option (Figure 2, fourth column from left, bottom row). These results suggest that girls from less advantaged social origins may be more likely to consider pragmatic compromises in terms of their first-ranked field's sex type compared to their peers with higher levels of parental resources. When extending the view to the next-best ranked options, the model suggests that such pragmatic compromises in terms of gender-atypical fields become more likely also for girls whose parents have medium-level education (Appendix Figure A.1, right column, bottom row).

Among boys, we find similar social gradients among first-ranked high school applications in the reference scenario, but a lower baseline share of boys who primarily apply to the academic track (Figure 3, first column from left, top row). Engineering is the most popular choice (Figure 3, first column from left, middle row), whereas the female-dominated health and social welfare field is decidedly unpopular among male applicants regardless of social origin (Figure 3, first column from left, bottom row).

In line with our binary models, the results in Figure 3 show boys with medium- rather than low-educated parents as more likely to adapt their application patterns to different distance constellations of upper secondary alternatives. But by differentiating upper secondary distances by field, our mother logit model results provide a slightly more nuanced angle to this finding. If high schools are the closest alternative, boys from families with medium levels of education (upper secondary or short tertiary qualifications) are predicted to mildly increase their share of first-ranked high school applications, but more clearly than boys with either low- or highly educated parents (Figure 3, second column from left, top row). On the other hand, if high schools are further away than other key options, the model predicts the clearest repercussions in the application patterns of boys with short-tertiary and highly-educated parents. In this scenario, first-ranked high school applications are predicted to decrease in these groups by about 5 pp, equivalent to a reduction of roughly 9-11% in relative terms (Figure 3, third and fourth column from left, top row). In terms of substitutions, (gender-typical) engineering absorbs the largest part of these applicants if engineering programmes are the closest alternative. However, in the scenario where (gender-atypical) health programmes are more easily accessible than other alternatives, our model predicts no noteworthy increase in boys applying to (gender-atypical) health and social welfare programmes (Figure 3, fourth column from left, bottom row) among boys (against expectation 4).

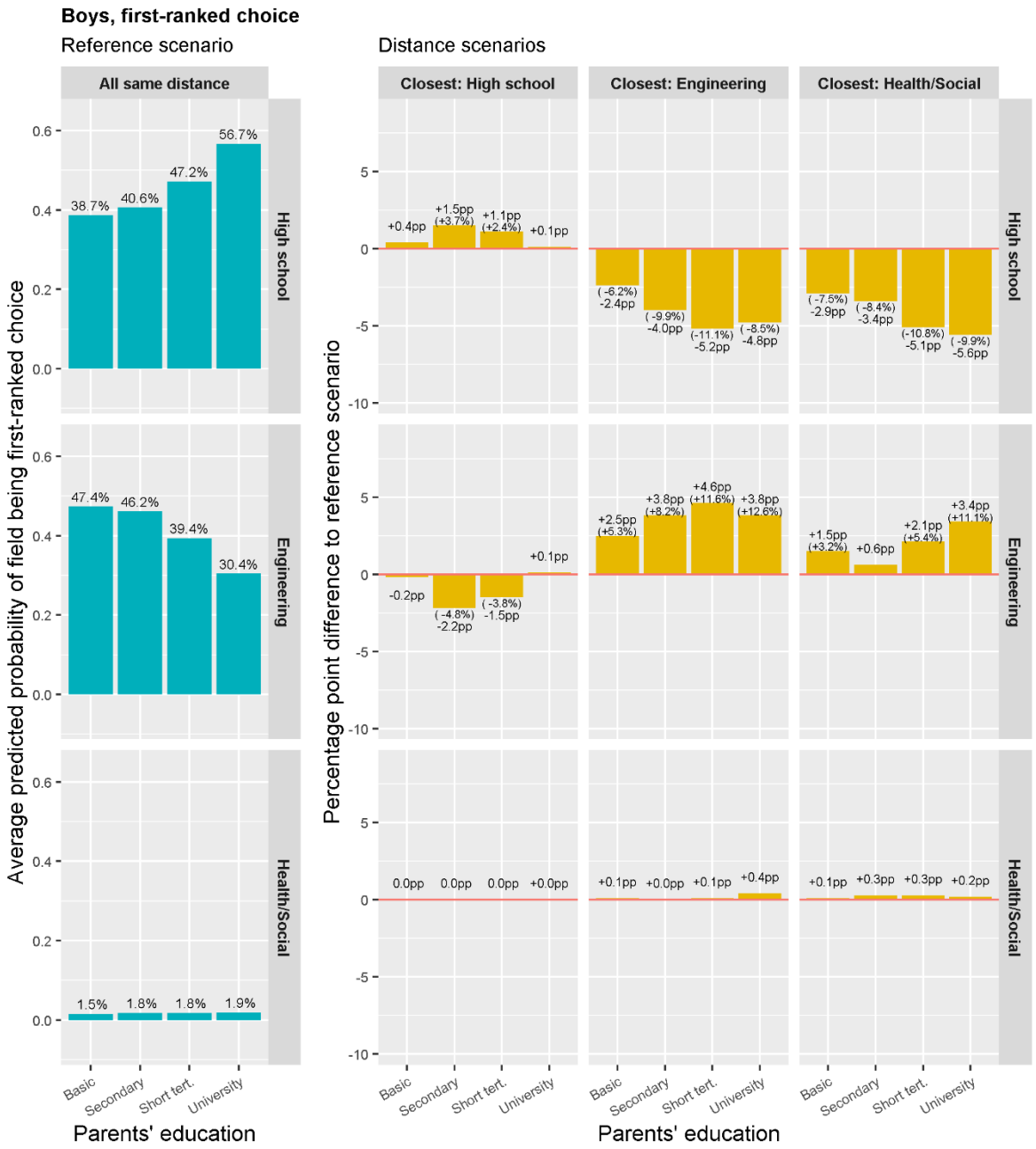


Figure 3. Boys' first-ranked choice by parental education and educational distance scenario. Average predicted probabilities in three hypothetical scenarios compared to a reference scenario (leftmost column). Results based on mother logit models with a three- distance×alternative×parental education interaction. For included predictors and goodness of fit statistics, see Appendix Table 1 (MoL model 2, boys).

Distance appears to matter also for boys selecting the next-best programme on their application (Appendix Figure A.2). For those that have not opted for a high school programme as their first choice, longer travel times to high schools appear to further discourage high school applications, regardless of social origin (Appendix Figure A.2, third and fourth column). More geographically accessible options now strengthen their role as substitutes. Importantly, if the health and welfare field is in closer vicinity than key alternatives, our model predicts boys with low-educated parents to be slightly more likely than their peers to consider such gender-atypical programmes as their next-

best choice (Appendix Figure A.2, fourth column from left, bottom row). This effect remains however very small in absolute terms (less than 1 pp).

5.3. Robustness checks

Finland is not only sparsely, but also very unevenly populated, which also means that distances to education can vary substantially between different regions (see also Table 2). To examine how sensitive the model averages reported in this study are to regional variation, we performed two types of robustness checks (see online appendix). First, we excluded the most densely populated Helsinki-Uusimaa region from the analyses of first-ranked track applications. Both the binary logit models as well as the mother logit analyses produced equivalent social origin gradients and substitution patterns, suggesting that the capital region did not substantially drive our original results. Second, we estimated separate models of first-ranked track and field/track combinations for each of the four NUTS2 regions of mainland Finland (Helsinki-Uusimaa, Southern Finland, Western Finland, Northern and Eastern Finland).

Among girls, social origin gradients in our binary logit model and overall substitution patterns of high schools for gender-(a)typical alternatives in our mother logit analyses were substantially equivalent in Helsinki-Uusimaa, Southern Finland and Western Finland. However, we found a flattening of social origin gradients when restricting our binary logit and mother logit models to female applicants residing in the North/East of the country. In Southern Finland, the (more gender-balanced) business field appeared to play a more important role in substituting distant high schools than the (female-typical) health and social care field. Among boys, the social origin gradients and substitution patterns we found in our models for the South, West, and the North/East of the country were substantively equivalent to our original results. However, we found a more uniformly negative effect of high school distances for boys from all social origins in our binary logit models of the Helsinki-Uusimaa region, which reappeared also in our mother logit models. The lack of gender-atypical substitutions and a general pattern of gender-typical substitutions for absent high schools replicated in all regional models among boys.

To summarize, our results remained for the most part robust across different regions of Finland, although a flattening of the social origin gradients for distance associations was noted for Helsinki (in the case of boys) and the North/East (in the case of girls).

6. Discussion and conclusion

Geographical accessibility is an increasingly recognized dimension of educational inequality in the sociological literature, although much research in this field has concentrated on tertiary education. In this article, we examined whether geographical distance and proximity might modify the salience of social constraints and opportunities for young people's upper secondary applications, by which we contributed a new angle on common explanatory accounts of social and gender inequalities in education. The rich, large-scale data on application patterns available to our study, linked with geographical distance measures of educational options, enabled us to observe young people's

educational plans before any institutional selection processes took place. As such, we were also better able to accommodate the diversity of educational alternatives that young people consider at this stage, of which we differentiated their first- and next-highest-ranked alternatives.

Our results showed that on average, geographical accessibility of education matters for young people's upper secondary pathways. In deciding their first-ranked upper secondary choice, we found young people to be more sensitive to the geographical distance of high schools, whereas travel times to the nearest vocational school appeared less significant for their choice of track. This suggests that applicants treat educational distances asymmetrically. Since applications to vocational schools imply the decision for a particular, fairly occupation-specific programme, the lower sensitivity to their location suggests that geographical proximity alone may not compensate for a possibly otherwise undesirable vocational specialization. High schools, on the other hand, perhaps due to their emphasis on general education, appear to be more likely to sway some applicants for more pragmatic reasons, if they are located closer to young people's home areas. We caution against misinterpreting this finding as implying that travel times to vocational schools do not matter to young people. Although distance may not discourage young people from applying to vocational schools, previous research has shown that longer travel times increase young people's risk of dropping out of upper secondary education (Falch et al., 2013; Virtanen & Riukula, 2021).

The question as to whether proximity incentivizes pragmatism is also relevant from a perspective of understanding social origin differences in educational pathways. In line with expectations derived from risk aversion accounts, we found that particularly daughters of highly educated parents were comparatively less sensitive in their upper secondary track choices to where high schools are located. For girls whose parents had received lower qualifications than a master's degree, on the other hand, we found stronger associations between upper secondary track choices and the geographical distance of educational alternatives. These social origin differentials in distance sensitivity are more in line with risk aversion accounts than with culture-based arguments. Rather than being systematically alienated from the educational system, daughters of parents with lower "cultural capital" in our study showed increased interest in academic pathways if geographical accessibility made such options more feasible. This finding aligns with the idea that what separates children of different backgrounds on their pathway through the educational system may be less a matter of class-specific culture than differences in uncertainty and risk, which geographical opportunity may to some (limited) extent either alleviate or aggravate.

However, the decidedly gendered ways in which social gradients of distance played out in our study also complicate an all too straightforward risk aversion interpretation. Daughters of parents with shorter qualifications were indeed most likely to adjust their upper secondary plans to the proximity or distance of high school options. However, boys from low- and highly-educated parents were less swayed by the distance or proximity of high schools than we expected. Instead, it was sons of parents with short-tertiary qualifications who were most likely to adjust their first preference to the geographical accessibility of high schools. In other words, among male applicants from lower-educated families, thresholds to consider high school pathways appear to be less malleable by structural opportunities arising through proximity.

Further nuances in these gender differences emerged once we extended our analyses to compare relative distance constellations of track- and field-specific educational alternatives. In (counterfactual) scenarios of different field- and track-specific distance constellations, our results predicted particularly girls from lower-educated social backgrounds as slightly more prepared to consider fields of study not typical for their gender. High levels of (parental) education may thus not in itself weaken young people's gender-stereotypical interests, and neither does the defiance of gender-stereotypical expectations appear to necessarily require high levels of cultural capital. It may be the case that challenges to gender-normative expectations are not exclusively rooted in progressive gender ideology, but can also emerge as a pragmatic adaption to constraints (Usdansky, 2011). Boys, on the other hand, substituted the absence of easily accessible high schools mainly via gender-typical vocational options, even if these were similarly distant as high schools. More rigid gender norms for men than women may be one aspect contributing to this finding (Bosson et al., 2022), which may lead to somewhat greater pragmatic leeway for girls rather than boys. Overall, however, some caution must be applied in interpreting the gender-(a)typical substitutions found in this study, as the rather broadly defined fields of study may also contain internal gender segregation.

To what extent distance to upper secondary education truly alters the risk and uncertainty young people associate with different upper secondary tracks cannot be definitively established within the context of this research. Furthermore, it is possible that institutional features differentially affect boys and girls as they consider their next steps on their way through the educational system. In systems like the Finnish one, where critical educational turning points coincide with pupils' late adolescence, previous studies have found more pronounced sex differences in attainment and upper secondary track choice, although it is unclear whether sex differences in (non-)cognitive development play a role for explaining this pattern (Scheeren & Bol, 2022). If the timing of upper secondary transitions falls into a "critical" period for development, this may also affect social origin and sex differences in the response to geographical constraints. Furthermore, as dead ends within the Finnish educational system are increasingly dismantled (exemplified by widespread possibilities to integrate high school subjects into a vocational qualification as a "double degree", or *kaksoistutkinto*), pressures to compromise between "safe" vocational options and future academic pathways may subside even further. If (gender-typical) generally more predictable vocational pathways do not necessarily come at the cost of excluding prospects for academic upward mobility aspirations, particularly boys from lower-educated social origins may simply not perceive the proximity of high schools as sufficiently salient for re-assessing their educational options. Further research into how constraints may translate into (gender-(a)typical) compromises is needed.

Finally, it is important to acknowledge several limitations characterizing this study. Although the administrative nature of this study's data helped minimize attrition in our study, the fact that young people's home address information was available only at the level of postcodes introduced some measurement error and reduced variation in our calculation of track- and field-specific distances to education. As such, distance calculations remained approximate, particularly for the country's more sparsely populated north and east. Furthermore, while our mother logit models were able to circumvent the problem of unrealistic modelling assumptions (such as the IIA) underlying conditional and multinomial logit models, some of the scenarios we compared in our analyses are empirically relatively rare in Finland. In this sense, our results included what could be considered out-of-sample predictions, which future research needs to subject to further scrutiny using different types of

methods and data. While our empirical analyses cannot determine the final causality of observed relationships in any definitive way, we expect the empirical and theoretical contributions of our study will assist future work with a more dedicated causal design.

7. References

Armila, P., Käyhkö, M., & Pöysä, V. (2018). On the educational edges of a learning society: The Finnish hinterland as a framework of educational choices for young people. *Journal of Youth Studies*, *21*(9), 1198–1215.

Barone, C., & Assirelli, G. (2020). Gender segregation in higher education: An empirical test of seven explanations. *Higher Education*, *79*(1), 55–78.

Barone, C., Barg, K., & Ichou, M. (2021). Relative risk aversion models: How plausible are their assumptions? *Rationality and Society*, *33*(2), 143–175.

Bertrand-Cloodt, D., Cörvers, F., Heijke, H., & Van Thor, J. (2010). The impact of distance deterrence on the choice of field of study in vocational education in the Netherlands. *Investigaciones de Economía de La Educacíon, Asociación de Economía de La Educación*, 605–624.

Bock, B. B. (2017). Gender and mobility. In B. B. Bock & S. Shortall (Eds.), *Gender and Rural Globalization: International Perspectives on Gender and Rural Development* (pp. 8–18). CABI.

Bosson, J. K., Wilkerson, M., Kosakowska-Berezecka, N., Jurek, P., & Olech, M. (2022). Harder won and easier lost? Testing the double standard in gender rules in 62 countries. *Sex Roles*, *87*(1–2), 1–19.

Bourdieu, P., & Passeron, J.-C. (1990). *Reproduction in Education, Society and Culture*. Sage.

Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials: Towards a formal rational action theory. *Rationality and Society*, *9*(3), 275–305.

Cedefop. (2015). *Finland—Reforming vocational qualifications*. European Centre for the Development of Vocational Training. https://www.cedefop.europa.eu/en/news-and-press/news/finland-reforming-vocational-qualifications

Chesters, J. (2021). Gender attitudes and occupational aspirations in Germany: Are young men prepared for the jobs of the future? *Work, Employment and Society*. https://doi.org/10.1177/09500170211017046

Denzler, S., & Wolter, S. C. (2011). Too far to go? Does distance determine study choices? *IZA Discussion Paper*, *5712*. https://ssrn.com/abstract=1849467

Dickerson, A., & McIntosh, S. (2013). The impact of distance to nearest education institution on the post-compulsory education participation decision. *Urban Studies*, *50*(4), 742–758.

Faggian, A., McCann, P., & Sheppard, S. (2007). Some evidence that women are more mobile than men: Gender differences in UK graduate migration behavior. *Journal of Regional Science*, *47*(3), 517–539.

Falch, T., Lujala, P., & Strøm, B. (2013). Geographical constraints and educational attainment. *Regional Science and Urban Economics*, *43*(1), 164–176.

Finger, C. (2016). Institutional constraints and the translation of college aspirations into intentions—Evidence from a factorial survey. *Research in Social Stratification and Mobility*, *46*, 112–128.

Flannery, D., & Cullinan, J. (2014). Where they go, what they do and why it matters: The importance of geographic accessibility and social class for decisions relating to higher education institution type, degree level and field of study. *Applied Economics*, *46*(24), 2952–2965.

Frenette, M. (2006). Too far to go on? Distance to school and university participation. *Education Economics*, *14*(1), 31–58.

Gibbons, S., & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional Science and Urban Economics*, *42*(1–2), 98–113.

Goldthorpe, J. H. (1996). Class analysis and the reorientation of class theory: The case of persisting differentials in educational attainment. *British Journal of Sociology*, *47*(3), 481–505.

Goldthorpe, J. H. (2007). Outline of a Theory of Mobility. In *On Sociology. Volume 2: Illustration and Retrospect* (pp. 154–185). Stanford University Press.

Gottfredson, L. S., & Lapan, R. T. (1997). Assessing gender-based circumscription of occupational aspirations. *Journal of Career Assessment*, *5*(4), 419–441.

Hällsten, M. (2010). The Structure of Educational Decision Making and Consequences for Inequality: A Swedish Test Case. *American Journal of Sociology*, *116*(3), 806–854. https://doi.org/10.1086/655751

Hällsten, M. (2017). Is education a risky investment? The scarring effect of university dropout in Sweden. *European Sociological Review*, *33*(2), 169–181.

Hango, D., Zarifa, D., Pizarro Milian, R., & Seward, B. (2019). Roots and STEMS? Examining field of study choices among northern and rural youth in Canada. *Studies in Higher Education*, 1–31.

Härtull, C., & Saarela, J. (2021). Educational mobility across three generations of Finnish-speaking and Swedish-speaking Finns. *Research in Social Stratification and Mobility*, *75*, 100649. https://doi.org/10.1016/j.rssm.2021.100649

Hess, S., Daly, A., & Batley, R. (2018). Revisiting consistency with random utility maximisation: Theory and implications for practical work. *Theory and Decision*, *84*(2), 181–204.

Huber, S., & Rust, C. (2016). Calculate travel time and distance with OpenStreetMap data using the Open Source Routing Machine (OSRM). *The Stata Journal*, *16*(2), 416–423.

Jacobs, J. A. (1989). *Revolving Doors: Sex Segregation and Women's Careers*. Stanford University Press.

Kilpi-Jakonen, E. (2011). Continuation to upper secondary education in Finland: Children of immigrants and the majority compared. *Acta Sociologica*, *54*(1), 77–106. https://doi.org/10.1177/0001699310392604

Knutsen, T. K., Modalsli, J., & Rønning, M. (2020). Distance and choice of field: Evidence from a Norwegian college expansion reform. *Statistics Norway Discussion Papers*, *932*.

Koskela, M. (2020). *Geocodes and addresses of high schools and field-specific vocational programmes.* Unpublished database.

Kozyreva, A., & Hertwig, R. (2021). The interpretation of uncertainty in ecological rationality. *Synthese*, *198*(2), 1517–1547.

Mize, T. D. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, *6*, 81–117.

Pigini, C., & Staffolani, S. (2016). Beyond participation: Do the cost and quality of higher education shape the enrollment composition? The case of Italy. *Higher Education*, *71*(1), 119–142.

Prix, I., & Kilpi-Jakonen, E. (2022). Not in a class of one's own: Social origin differentials in applying to gender-(a)typical fields of study across the educational hierarchy. *European Sociological Review*, *38*(6), 920–941. https://doi.org/10.1093/esr/jcac007

Pulkkinen, J., & Rautopuro, J. (2018). *Opin polut ja pientareet harvaan asutuilla alueilla. Nuorisobarometri 2017 lisäotos*. Valtion nuorisoneuvosto, OKM.

Rönnlund, M. (2020). 'I love this place, but I won't stay': Identification with place and imagined spatial futures among youth living in rural areas in Sweden. *Young*, *28*(2), 123–137.

Sá, C., Florax, R. J., & Rietveld, P. (2006). Does accessibility to higher education matter? Choice behaviour of high school graduates in the Netherlands. *Spatial Economic Analysis*, *1*(2), 155–174.

Scheeren, L., & Bol, T. (2022). Gender inequality in educational performance over the school career: The role of tracking. *Research in Social Stratification and Mobility*, *77*, 100661. https://doi.org/10.1016/j.rssm.2021.100661

Seehuus, S. (2019). Social class background and gender-(a) typical choices of fields of study in higher education. *The British Journal of Sociology*, *70*(4), 1349–1373.

Spiess, C. K., & Wrohlich, K. (2010). Does distance determine who attends a university in Germany? *Economics of Education Review*, *29*(3), 470–479.

Statistics Finland. (2018). *Gender equality in Finland 2018*. Statistics Finland. https://tilastokeskus.finna.fi/Record/doria_tk.10024_184392?imgid=1

Statistics Finland. (2023). *Population and upper secondary application register data modules. User licence TK/27/07.03.00/2022.*

Steenburgh, T. J. (2008). The invariant proportion of substitution property (IPS) of discrete-choice models. *Marketing Science*, *27*(2), 300–307.

Stenström, M.-L., & Virolainen, M. (2014). *The current state and challenges of vocational education and training in Finland*. Nord-VET.

Suhonen, T. (2014). Field-of-study choice in higher education: Does distance matter? *Spatial Economic Analysis*, *9*(4), 355–375.

Thaning, M. (2021). Resource specificity in intergenerational inequality: The case of education, occupation, and income. *Research in Social Stratification and Mobility*, *75*, 100644. https://doi.org/10.1016/j.rssm.2021.100644

Turley, R. N. L. (2009). College proximity: Mapping access to opportunity. *Sociology of Education*, *82*(2), 126–146.

Tuuva-Hongisto, S., Pöysä, V., & Armila, P. (2016). Syrjäkylien nuoret–unohdetut kuntalaiset. *Helsinki: Kunnallisalan kehittämissäätiön, Tutkimusjulkaisu-sarjan julkaisu*, *99*.

Usdansky, M. L. (2011). The gender-equality paradox: Class and incongruity between work-family attitudes and behaviors. *Journal of Family Theory & Review*, *3*(3), 163–178.

Vipunen Education Statistics Finland. (2022). *Direct admission into further education by level of education (reference periods between 0 and 1 years)*. https://vipunen.fi/en-gb/_layouts/15/xlviewer.aspx?id=/en-gb/Reports/Peruskoulun%20lukion%20ja%20ammatillisen%20koulutuksen%20koulutukseen%20sijoittuneet%20-%20ammatillinen_EN.xlsb

Vipunen Education Statistics Finland. (2023). *Applicants and selected candidates*. https://vipunen.fi/en-gb/combined/Pages/Hakeneet-ja-hyv%C3%A4ksytyt.aspx

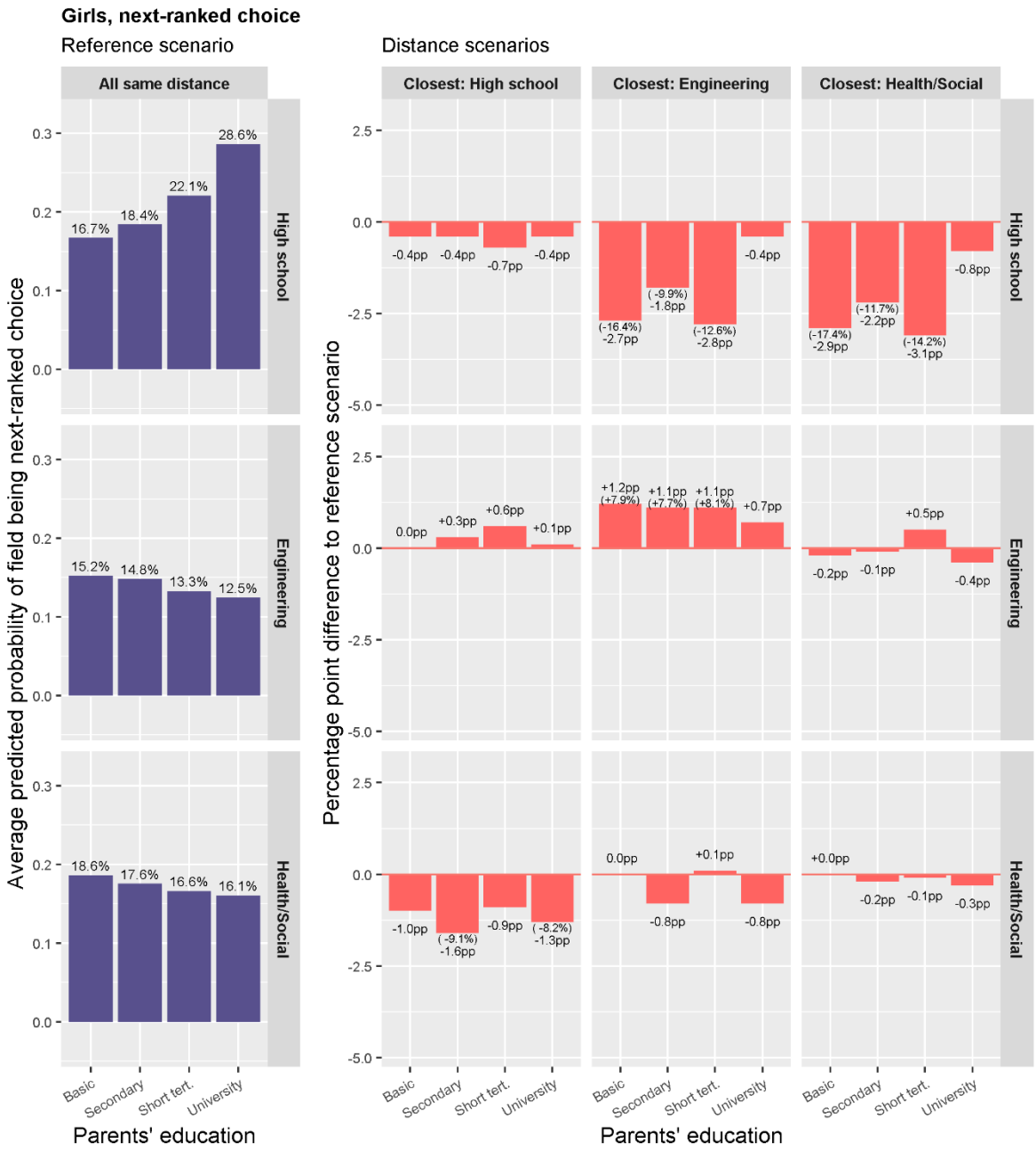
Virtanen, H., & Riukula, K. (2021). The effect of access to post-compulsory education: Evidence from structural breaks in school supply. *ETLA Working Papers*, *85*. https://www.etla.fi/wp-content/uploads/ETLA-Working-Papers-85.pdf

Appendix

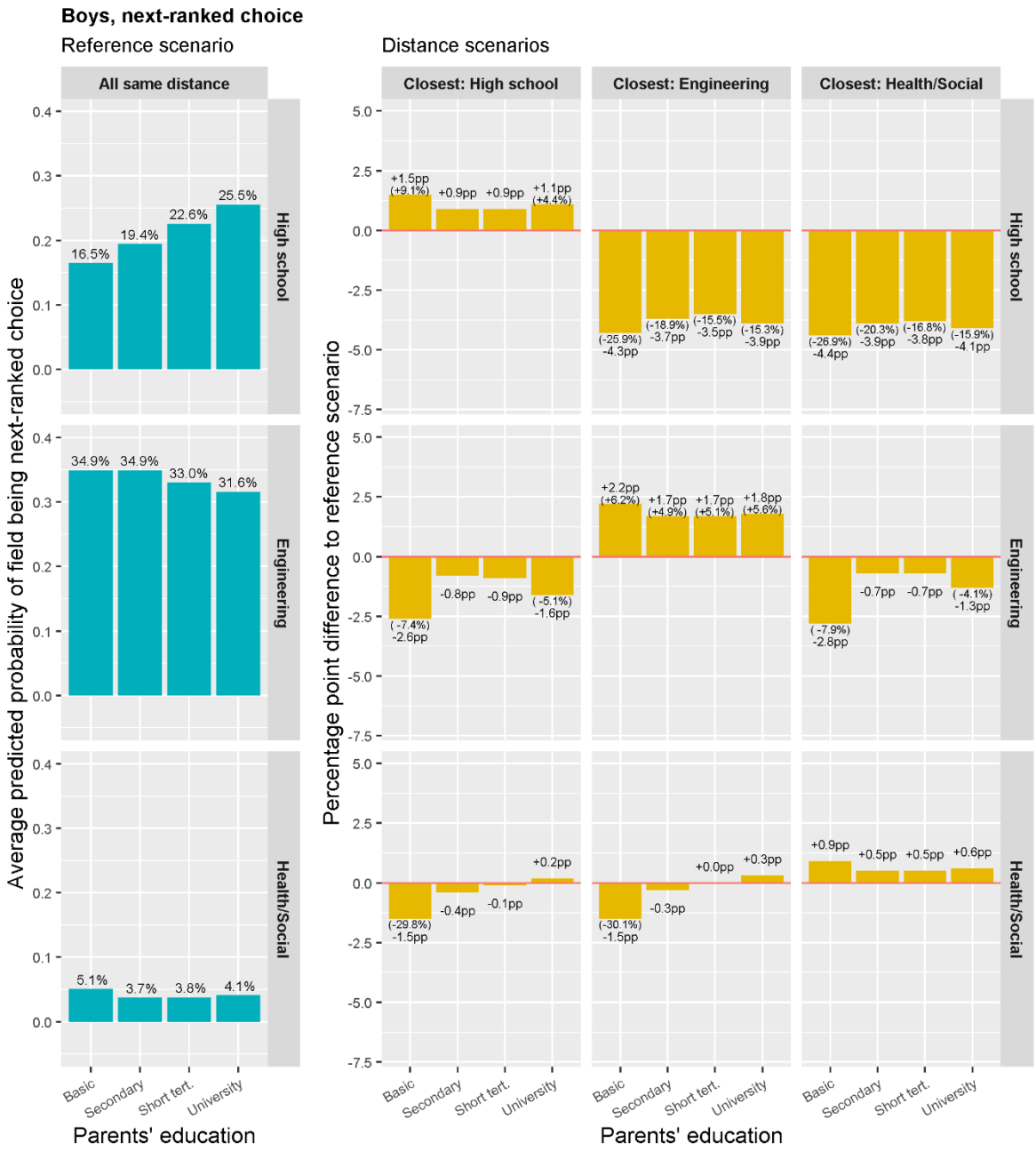
Appendix Table A.1. Comparison of goodness of fit indicators for conditional logit (CL) and mother logit models (MoL) of young people's first-ranked upper-secondary alternative.

	Boys				Girls			
	CL model 1	CL model 2	MoL model 1	MoL model 2	CL model 1	CL model 2	MoL model 1	MoL model 2
LR test (p-value)	ref.	73.12 (0.000)	246.12 (0.000)	242.46 (0.000)	ref.	38.60 (0.011)	126.48 (0.000)	122.09 (0.000)
AIC	235,542.4	235,511.3	235,356.3	235,328.9	283,213.8	283,217.2	283,147.3	283,155.1
N (person-alternatives)	954,457	954,457	954,457	954,457	955,472	955,472	955,472	955,472
N (individuals)	136,351	136,351	136,351	136,351	136,496	136,496	136,496	136,496

Note: CoL = Conditional logit model, MoL = Mother logit model. Model 1: includes distance*alternative interaction. Model 2: includes distance*alternative*parental education -interaction. All models include controls for parental education, equivalized household income (squared), GPA in lower secondary school, family type, mother tongue, NUTS2-region, urbanicity, regional field-specific unemployment rate, having immigrant parents.



Appendix Figure A.1. Girls' next-ranked educational choice on their upper secondary application in four hypothetical scenarios. Average predicted probabilities in three hypothetical scenarios compared to a reference scenario (leftmost column). Results based on mother logit models with distance×alternative×parental education interaction. Model specification identical with Figure 2.



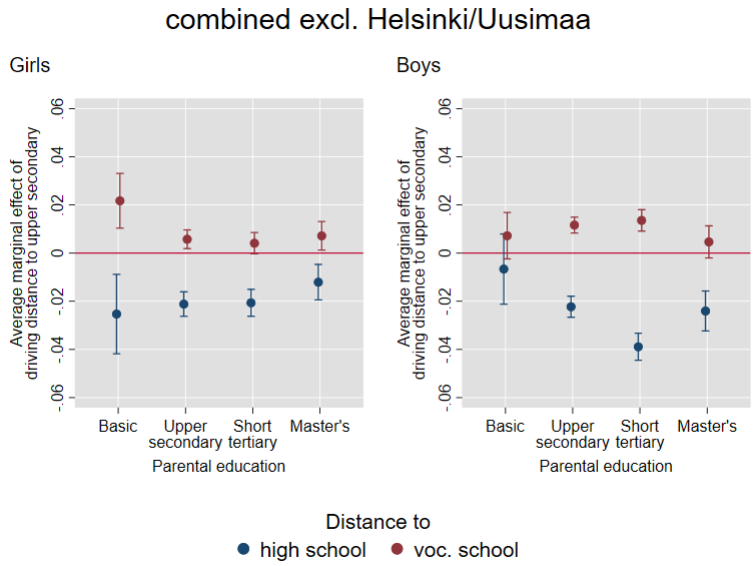
Appendix Figure A.2. Boys' next-ranked educational choice on their upper secondary application in four hypothetical scenarios. Average predicted probabilities in three hypothetical scenarios compared to a reference scenario (leftmost column). Results based on mother logit models with distance×alternative×parental education interaction. Model specification identical with Figure 3.

ONLINE APPENDIX

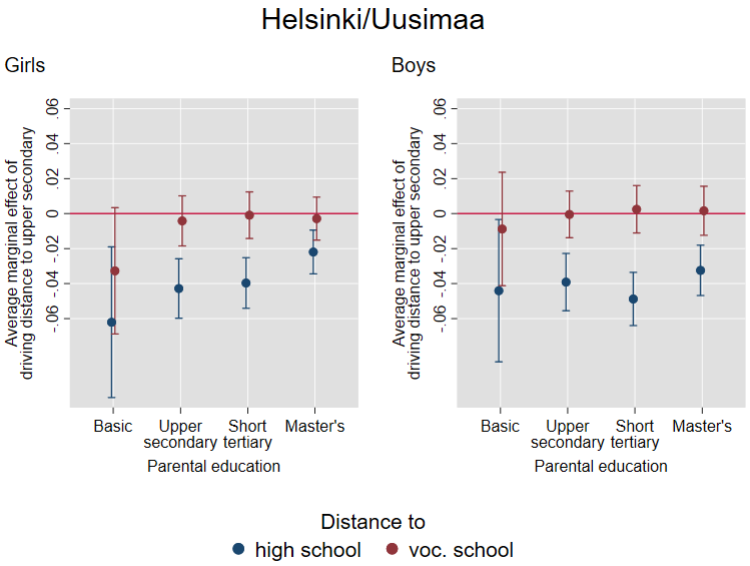
1. Region-specific interactions between social origin and upper secondary distances

The figures in this section plot average marginal effects of upper secondary distances on the probability of ranking high schools as the first choice, by parental background, gender and region of residence. Plotted average marginal effects are based on binary logit models with parental education×high school distance and parental education×vocational school distance –interactions, with models separately estimated by region of residence and gender. All models control for urbanization degree of home municipality, GPA in lower secondary school (squared), family type, average equivalized household income (squared), mother tongue, having immigrant parents.

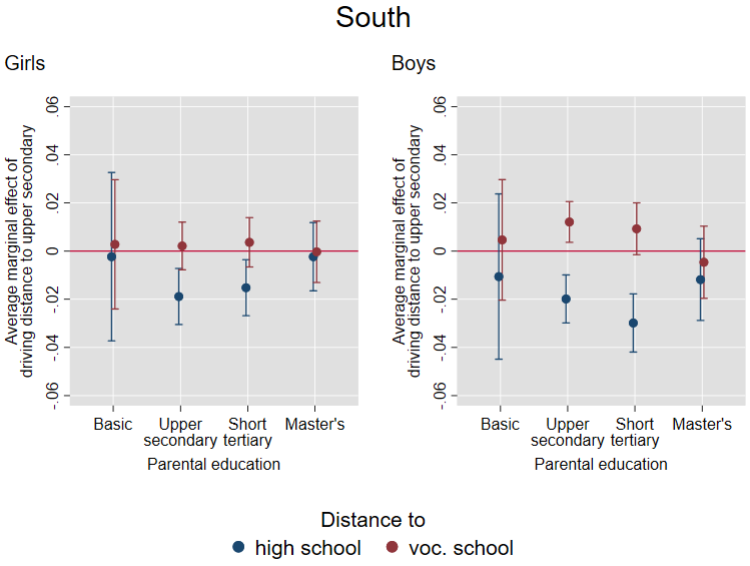
a. Combined models excluding the region Helsinki-Uusimaa



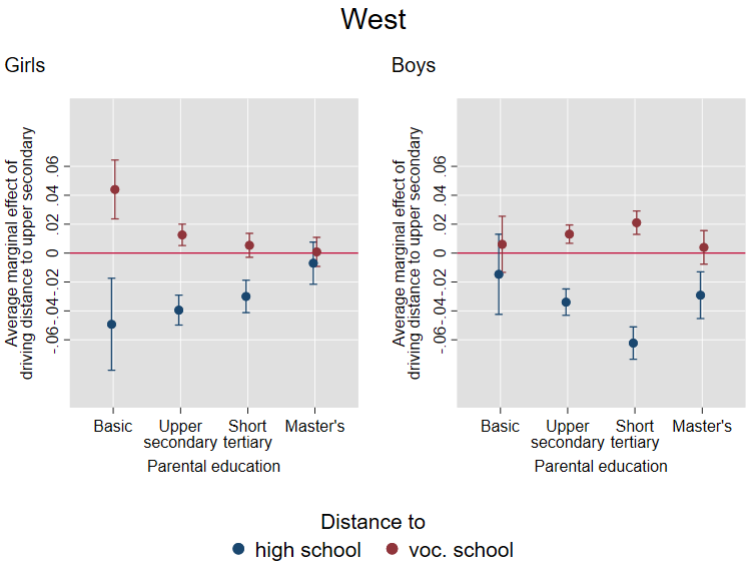
b. Region-specific models for Helsinki-Uusimaa



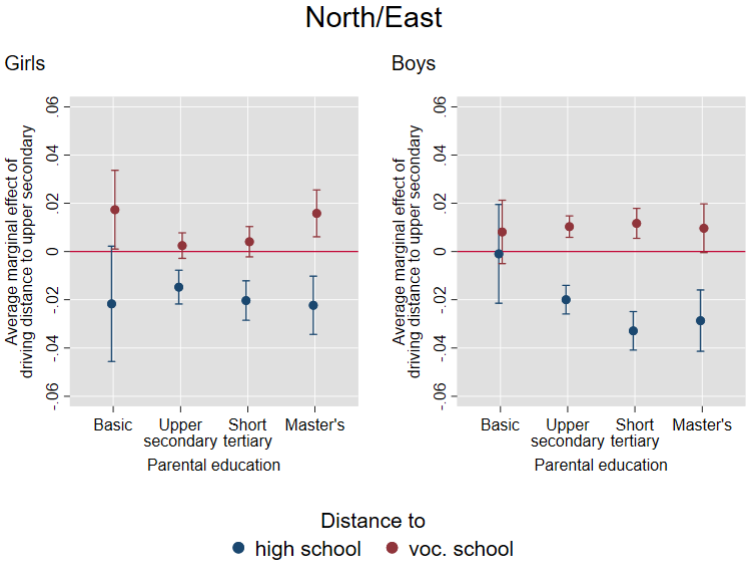
c. Region-specific model for Southern Finland



d. Region-specific model for Western Finland



e. Region-specific model for Northern and Eastern Finland

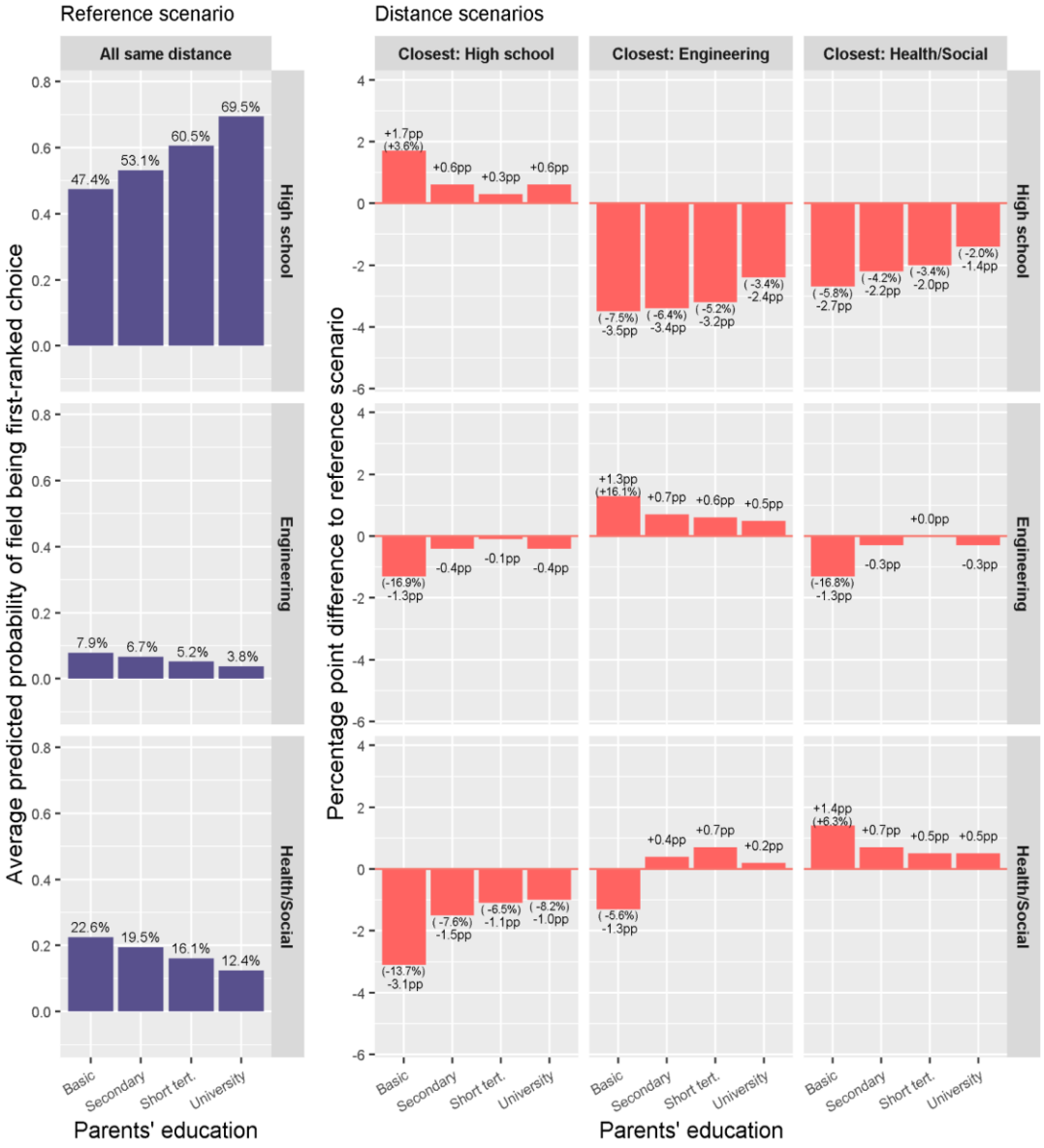


2. Region-specific predictions of substitution patterns in various distance constellations (based on region-specific mother logit models)

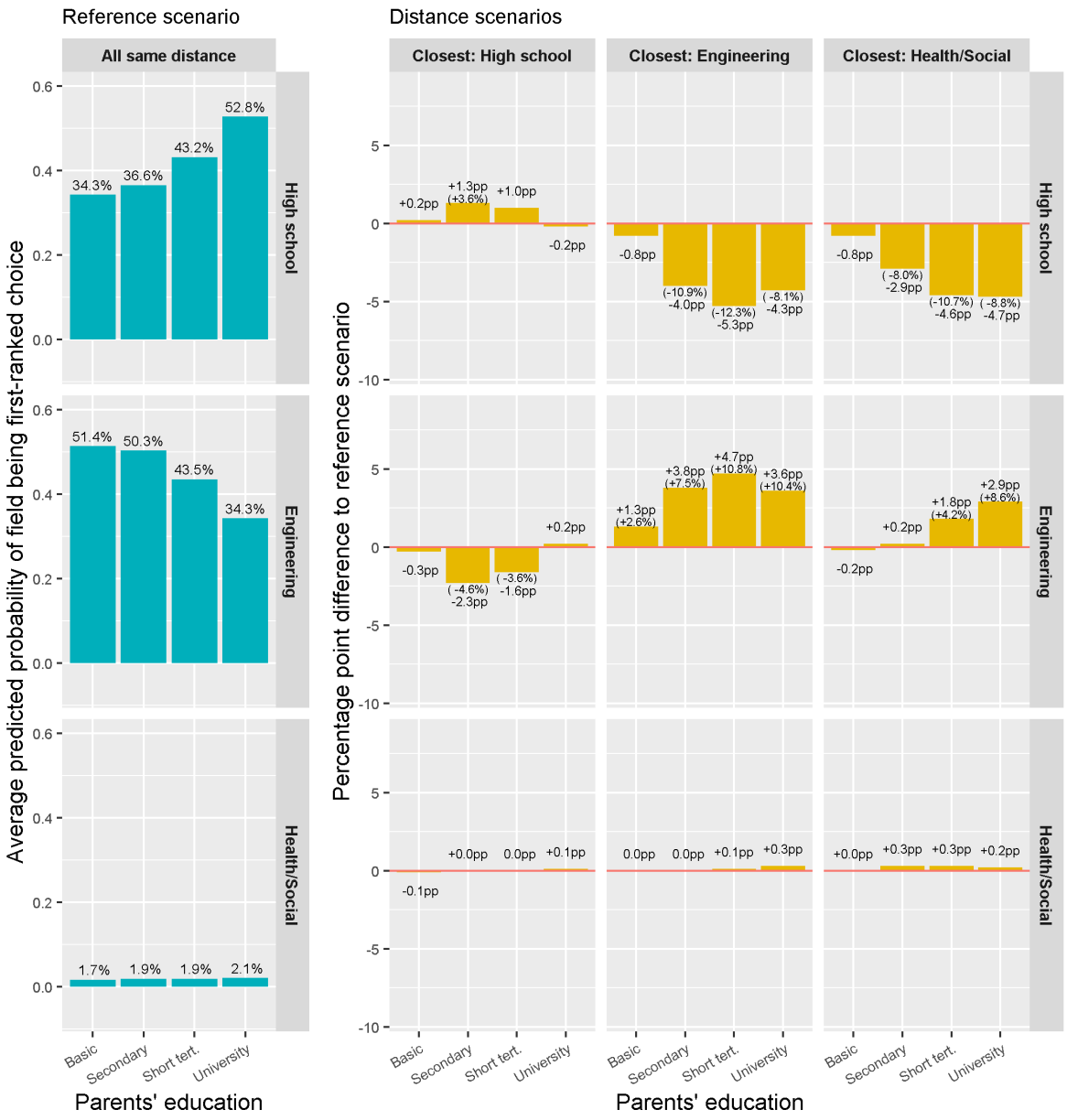
The figures in this section display average predicted probabilities of applicants' first-ranked choice by parental education and educational distance scenario. Average predicted probabilities in three hypothetical scenarios are compared to a reference scenario (leftmost column). Results are based on mother logit models with a three-way distance×alternative×parental education interaction. The figures in this section are based on models estimated separately for region of residence and gender. All models include controls for equivalized household income (squared), GPA in lower secondary school, family type, mother tongue, urbanicity, regional field-specific unemployment rate, having immigrant parents.

a. Combined model excluding Helsinki-Uusimaa

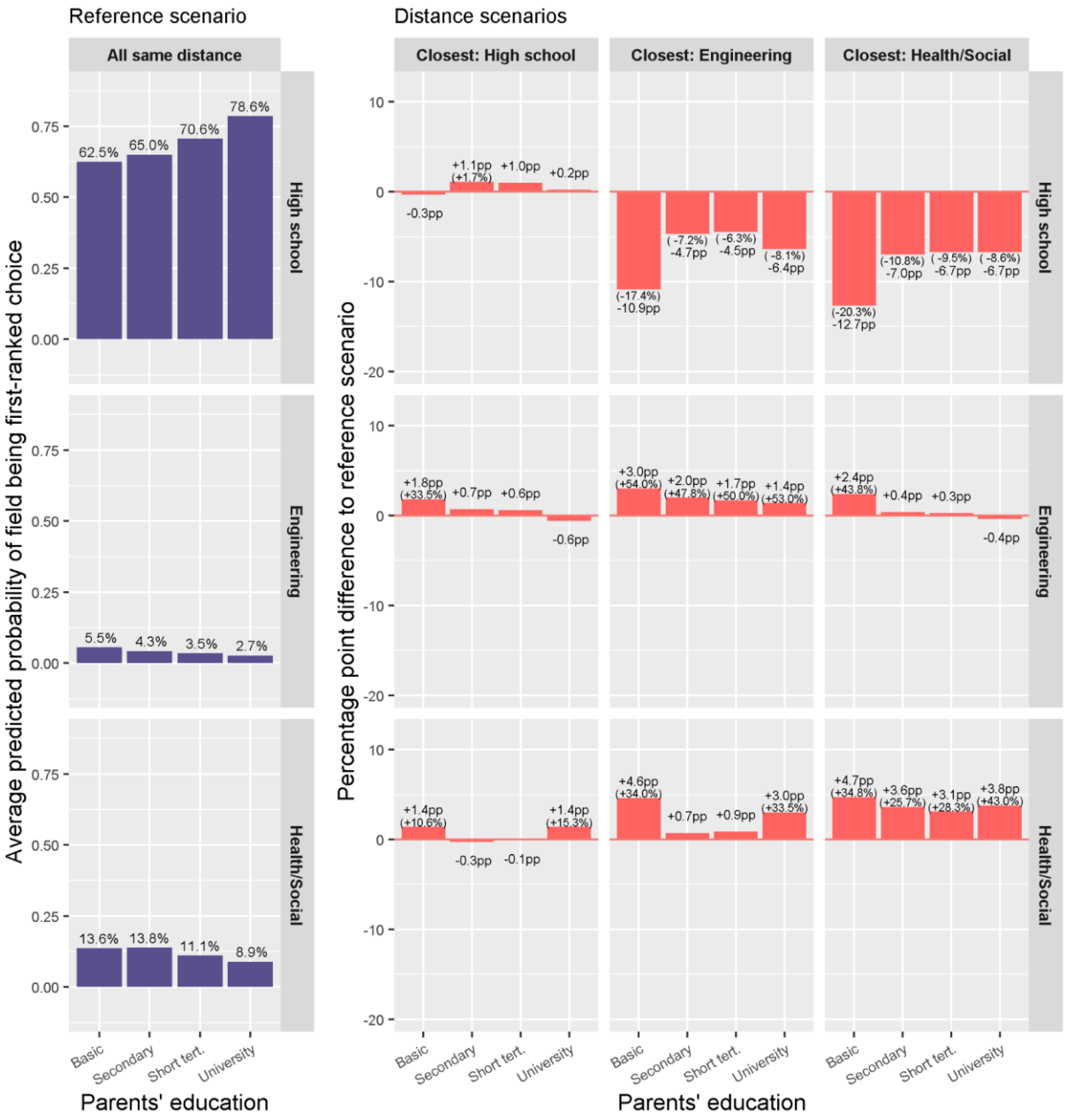
Girls, first-ranked choice: all regions excluding Helsinki/Uusimaa



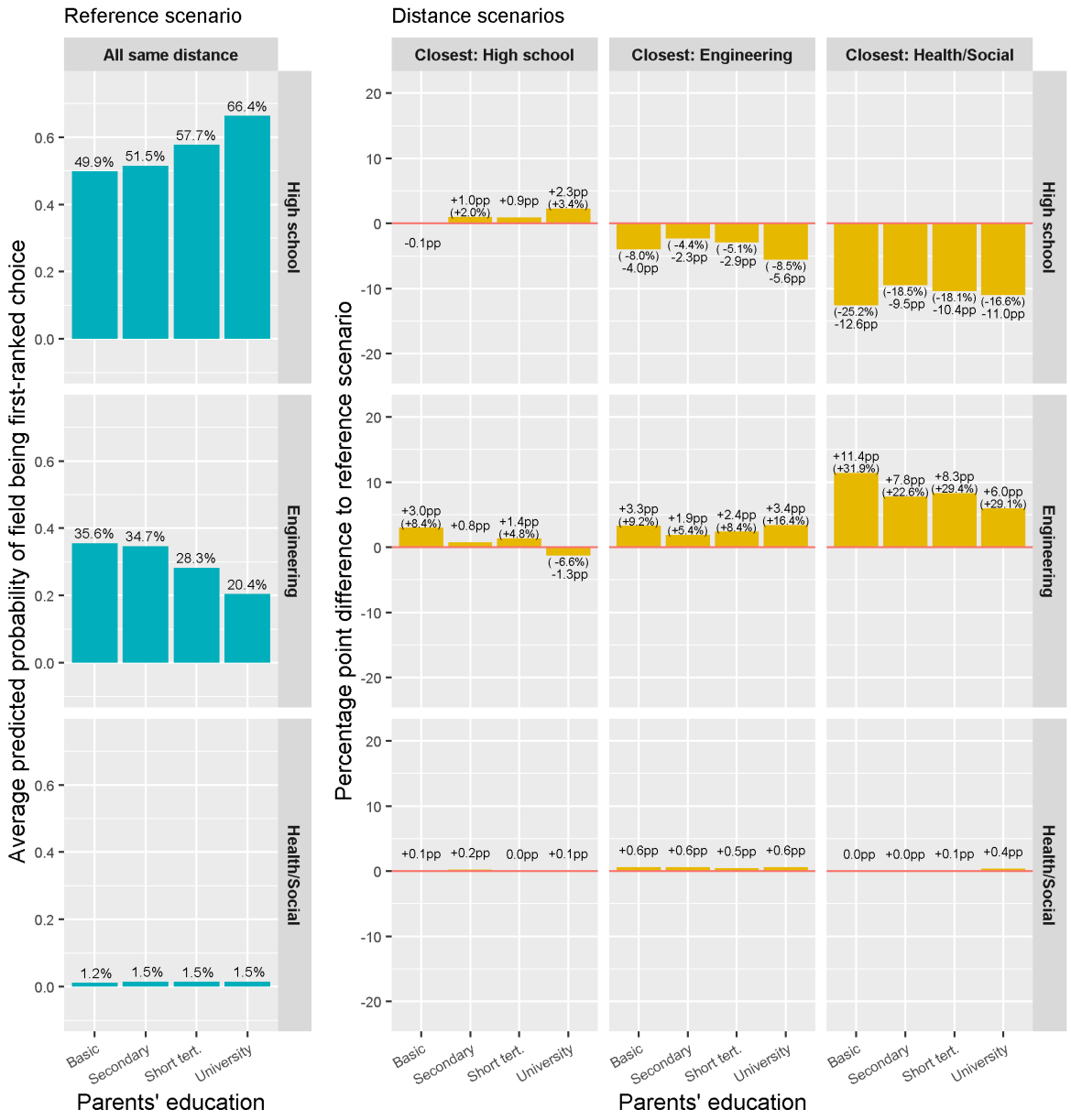
Boys, first-ranked choice: all regions excluding Helsinki/Uusimaa



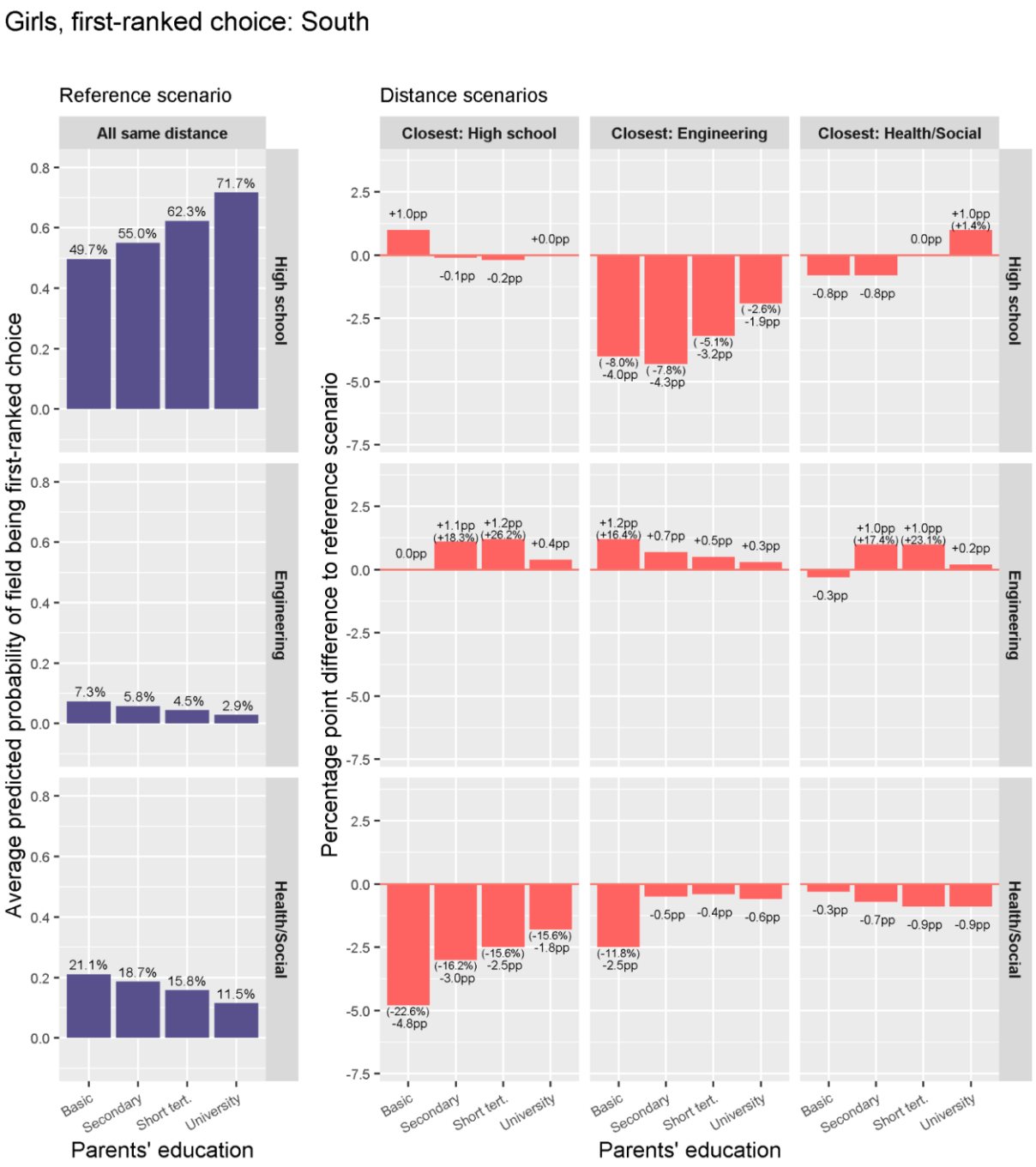
b. Region-specific model for Helsinki-Uusimaa



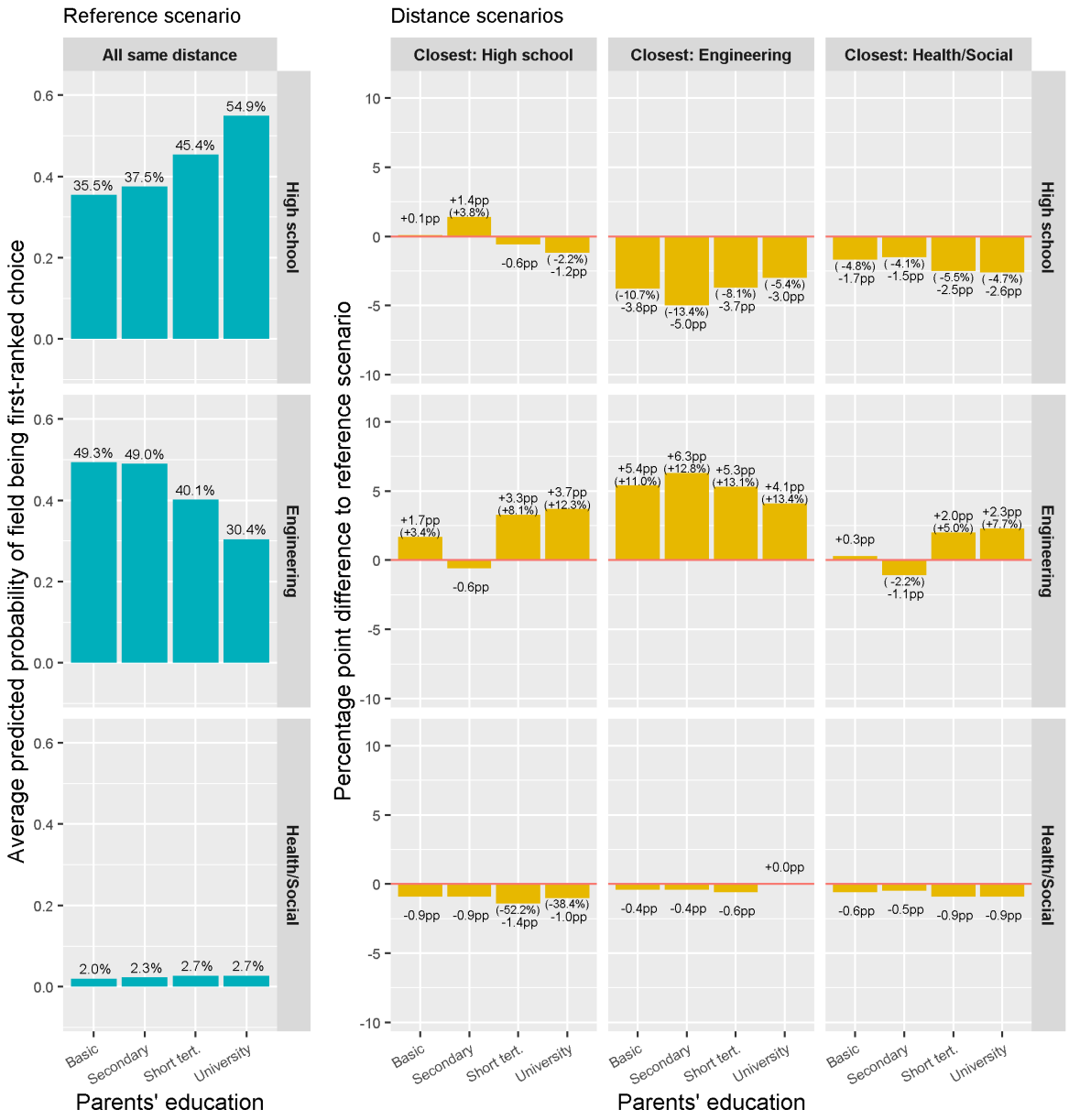
Boys, first-ranked choice: Helsinki/Uusimaa



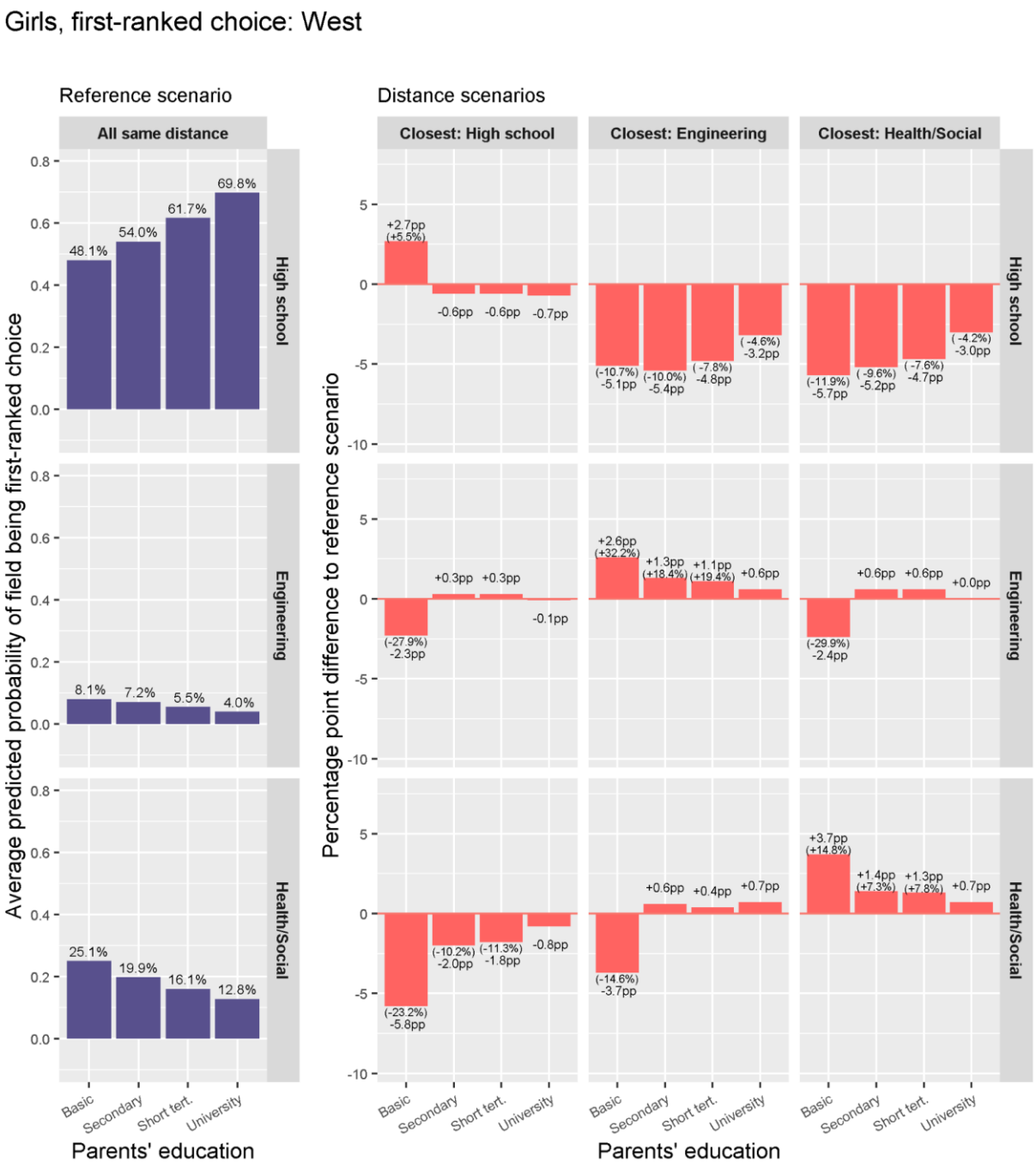
c. Region-specific model for Southern Finland



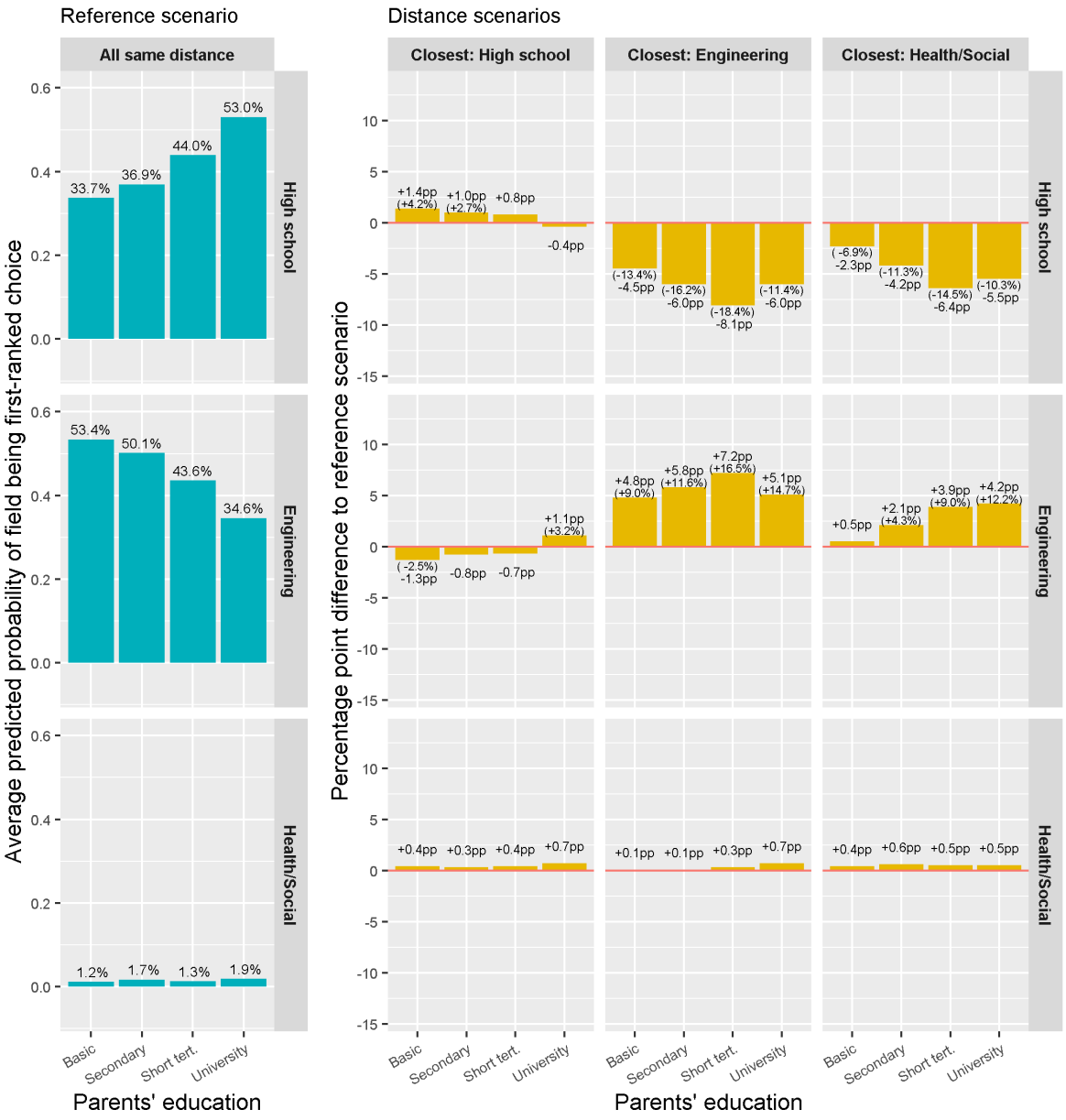
Boys, first-ranked choice: South



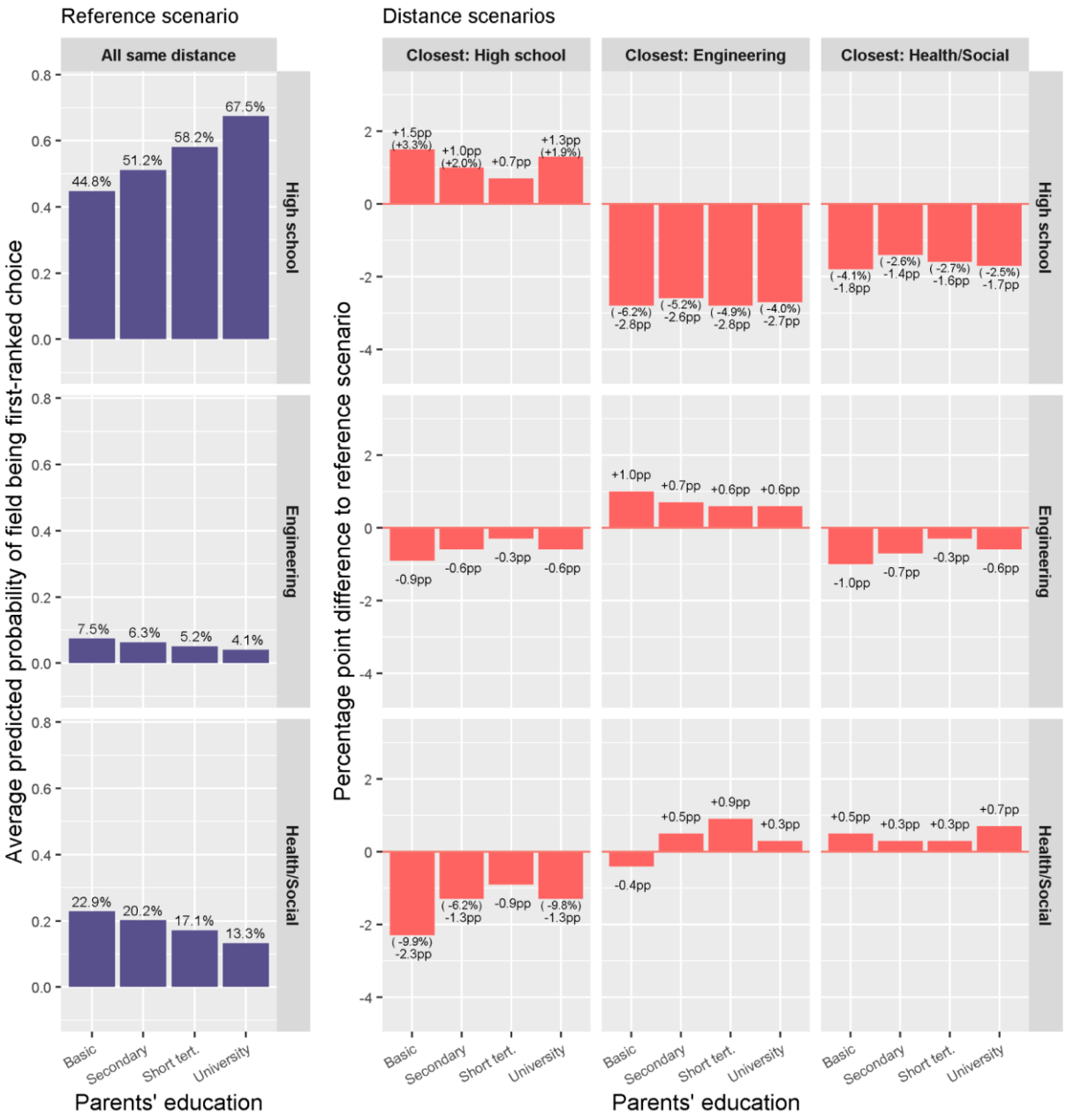
c. Region-specific model for Western Finland



Boys, first-ranked choice: West



d. Region-specific model for Northern and Eastern Finland



Boys, first-ranked choice: North/East